2012

Strategies for recombinant protein production in maize

Xing Xu
Iowa State University

Strategies for recombinant protein production in maize

by

Xing Xu

A dissertation submitted to the graduate faculty
In partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Plant Biology

Program of Study Committee:
Kan Wang, Major Professor
Charles Glatz
Martha James
Lawrence Johnson
Paul Scott

Iowa State University
Ames, Iowa
2012

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my major professor and mentor Dr. Kan Wang, for her support and patience throughout my PhD study. She has not only provided support for this research, but also trained me how to be a scientist. Without her help and encouragement, I would never have completed this work.

I would like to thank my committee members, Drs. Charles Glatz, Martha James, Lawrence Johnson, and Paul Scott for providing ideas and suggestions for this research. Special thanks to my collaborators including Julio Baez, Joan Cunnick, Qinglei Gan, Lorena Moeller, and Liyan Yang for their contributions and help.

I also would like to thank members of the Glatz lab including Matt Aspelund, Kerry Campbell, William Colonna, and Cheng Zhang for technical support and suggestions. Thanks to Bronwyn Frame, Marcy Main, Tina Paque, Rosemarie Schick and Katey Warnberg of Plant Transformation Facility for producing transgenic maize and providing tissue culture and greenhouse related assistances. Thanks to Margie Carter, Siquan Luo, and Joel Nott in the Protein Facility and Ann Perera in W.M. Keck Metabolomics Research Laboratory at Iowa State University for discussions and suggestions. I also thank Sarah Hosford, Angela Nguyen, and Jessica Zimmer for their technical assistance.

I appreciate the friendship I received from my colleagues and friends including Fan Cao, Jing Fang, Bronwyn Frame, Qinglei Gan, Yan Jing, Keunsub Lee, Diane Luth, Marcy Main, Susana Martin-Ortigosa, Ksenija Markovic, Juan Carlos Martinez-Nicolas, Lorena Moeller, Hartinio Nahampun, Tina Paque, Rose Schick, Andrea Scarpa, Francois Torney, Katey Warnberg, Chichun Yang, and Liyan Yang.

I especially thank my undergraduate advisor Huixia Shou for her support and help. Without her encouragement, I likely would not be embarking upon becoming a scientist.

This dissertation is dedicated to my mother, Zhiren Zhang, for her love, support and encouragement.

ABSTRACT

Maize is not only one of the world's most important crops, but it is also the first field-grown plant-based recombinant expression system developed for commercial purposes. With advantages of low cost, high yield, high protein stability, and well-developed transformation technology over other plant systems, maize is considered to be an ideal recombinant production system. However, there are three major challenges limiting the use of maize for producing recombinant proteins: 1) maize lacks mammalian-like post-translational modification ability that may be required for pharmaceutical and industrial proteins; 2) transgenic pollen containment in open-field production; and 3) low expression of recombinant products. The aim of the present study was to develop strategies to safely produce recombinant proteins with high yield and adequate modification in maize seeds.

Firstly, we showed that the recombinant human collagen type I alpha 1 (rCIα1) with high ratio of hydroxylated prolines (Hyp) similar to native human collagen was expressed in maize seeds co-expressing a recombinant human prolyl 4-hydroxylase (rP4H). The hydroxylated rCIα1 had a markedly enhanced thermal stability compared to non-hydroxylated rCIα1.

Secondly, we demonstrated that a cytoplasmic male-sterile maize system could be used for open-field production of maize-derived recombinant protein to minimize contamination risk caused by pollen drift. Two strategies were presented: the transgene was directly transformed into a male-sterile line via biolistic- or *Agrobacterium*-mediated methods, or introgressed from male-fertile to male-sterile germplasm by conventional breeding. The male-sterile transgenic maize seed from the second strategy can be used for

open-field production by using a non-transgenic maize pollen donor to produce 100% transgenic seeds after six seasons of breeding.

Thirdly, we examined the feasibility of producing a small peptide in maize seeds. A designed 22 amino acid peptide biosurfactant GAM1 was fused to the C-terminus of a fluoresce protein, the red fluorescent protein (RFP), or a maize seed storage protein, the 22-kDa α-zein, to avoid proteolytic degradation in the cell. While red fluorescence could be detected in seeds carrying RFP fusion construct, we failed to detect the presence of GAM1 despite using various protein concentration and detection methods. Strategies for producing this peptide in maize will need to be redesigned.

Lastly, we evaluated two strategies for enhancing recombinant protein expression in maize seeds. In the first strategy, we generated a line with reduced seed storage protein γ-zein by RNA interference (γRNAi). Then we crossed this line with a transgenic line expressing green fluorescent protein (γGFP). The GFP levels in subsequent generations were decreased in the γGFP/γRNAi seeds, which was opposite to what we expected. In the second strategy, we generated a line (oxPBF) constitutively overexpressing the transcription factor prolamin-box binding factor (PBF) that plays regulatory roles to the 27-kDa γ-zein promoter. We intend to cross the oxPBF line with transgenic lines expressing the GFP or recombinant proteins driven by the 27-kDa γ-zein promoter. It is hoped that the PBF overexpression may enhance the production of foreign protein under the γ-zein promoter control.

This work indicated that maize has the potential to produce recombinant proteins with mammalian-like post-translational modifications and can be used for open-field production when using a cytoplasmic male-sterile breeding system.

CHAPTER 1: INTRODUCTION

1.1. General Introduction

1.1.1. Brief history of plant molecular farming

Plants, ranging in size from several millimeters to more than one hundred meters, provide us with food, fuel, fiber, and the most important element required to sustain our life—oxygen. But, people have been trying to obtain more from plants. As a result of fast developing biotechnology during past few decades, nowadays plants can be used to produce various heterologous proteins, including pharmaceutical and industrial proteins, through recombinant DNA technology, often referred to as plant molecular farming (Faye and Gomord, 2010; Ma and Wang, 2012; Obembe et al., 2011; Wilken and Nikolov, 2012). In 1970, the first isolation of a restriction endonuclease, *Hin*d III, was reported (Smith and Wilcox, 1970). This study opened the door for recombinant DNA technology. The 1978 Nobel Prize for Physiology or Medicine was awarded for this discovery. The first recombinant DNA experiment was reported by Cohen and Boyer in 1973, twenty years after the discovery of the DNA structure by Watson and Crick (Cohen et al., 1973). They cut a fragment of DNA from the plasmid of one *E. coli* and transferred it into another *E. coli* DNA. Today, DNA is not only able to be transferred, but also can be artificially synthesized. Plus, with developing transformation techniques, the choice of the expression system is no longer restricted to bacteria. Yeast, mammalian cell culture, transgenic animals, plant cell cultures, and transgenic plants are common hosts used to produce foreign proteins. In this thesis, I present strategies for producing recombinant proteins in plants by using maize as a model system.

1.1.2. Advantages of plants over other recombinant expression systems

Plants are chosen because they are ideally suitable for producing recombinant proteins with many advantages over other systems (Ma et al., 2003). First, compared to the traditional microbial fermentation, plants are able to perform post-translational modifications, such as glycosylation and hydroxylation, which are required for biological activity of numerous mammalian proteins (Hood et al., 1997; Tremblay et al., 2010; Xu et al., 2011b). Second, plants, especially crops, can produce exogenous proteins on large scale at very low cost, since plant growth, transport, and post-harvest process requirements are relatively inexpensive (Fischer and Emans, 2000; Giddings, 2001). Third, unlike animal or microorganism systems, plant-derived products reduce contamination risk due to human pathogens (Ma et al., 2003; Streatfield et al., 2001). There are also drawbacks for transgenic plant production; one of the concerns is that the timescale of production is relatively long compared with plant or animal cell cultures and microorganisms (Ma et al., 2003).

1.1.3. Plant-derived products

After the first recombinant protein, human insulin, was approved as a drug by the United States Food and Drug Administration (FDA) in the 1980s, the demand for recombinant proteins, especially pharmaceutical proteins, dramatically increased (Ma and Wang, 2012). Almost at the same time, improved plant transformation techniques and strategies made the plant capable of being a recombinant protein expression system (Bevan et al., 1983). The first plant-derived recombinant pharmaceutical protein, a human growth hormone, was produced in transgenic tobacco in 1986 (Barta et al., 1986).

Not long after, the first plant-derived antibody was reported in 1989 (Hiatt et al., 1989). Then, the first plant-derived vaccine, hepatitis B virus surface antigen, was expressed in transgenic tobacco in 1992 (Mason et al., 1992). In 1997, the maize-derived recombinant avidin, an egg protein, became the first plant-derived recombinant protein for commercial purposes (Hood et al., 1997). The plant-derived recombinant products have now been extended to a broader area, which can be generally categorized into three groups (Xu et al., 2011a): 1) therapeutic proteins, such as antibodies (Lai et al., 2010; Ma et al., 1998), blood proteins (Sijmons et al., 1990), cytokines (Liu et al., 2009), growth factors (Parsons et al., 2010), growth hormones (Staub et al., 2000), and vaccines (Chikwamba et al., 2002a; Chikwamba et al., 2002b; Shoji et al., 2008; Shoji et al., 2011; Takaiwa, 2009); 2) industrial enzymes, such as avidins (Hood et al., 1997; Murray et al., 2002), α-amylase (Pen et al., 1992), β-glucuronidase (Kusnadi et al., 1998), cellulase (Hood et al., 2012; Hood et al., 2007), laccase (Hood et al., 2003), and trypsin (Woodard et al., 2003); and 3) biopolymers, such as collagens (Ruggiero et al., 2000; Xu et al., 2011b; Zhang et al., 2009a; Zhang et al., 2009b), elastin-like polypeptides (Conley et al., 2009), plant gum (Xu et al., 2005), and spider silk proteins (Menassa et al., 2004; Scheller et al., 2001; Yang et al., 2005).

1.1.4. Seed-based expression

Recombinant proteins are commonly expressed in plant seeds, including many cereal grains like barley, maize, rice, soybean, and wheat (Ramessar et al., 2008). Compared with leaf tissue, seeds are better for protein accumulation and storage (Boothe et al., 2010; Lau and Sun, 2009; Ma et al., 2003). Seeds have advantages of in-field

process and long-term storage. High expression of recombinant proteins has been achieved by using seed-targeting strategies up to 10% of total seed proteins (Clough et al., 2006; Hood et al., 2007; Hood et al., 1997; Horvath et al., 2000; Woodard et al., 2003; Xie et al., 2008; Xue et al., 2003; Yang et al., 2006; Yang et al., 2008).

1.1.5. Maize seed as an ideal production platform

Maize is one of the world's major crops, domesticated from teosinte more than 5,000 years ago (Gewin, 2003). The United States produces approximately 40% of the world's maize. In 2011, the production of maize in the U.S. was 12 billion bushels (http://www.nass.usda.gov/). Maize not only provides food and feed, but is also an important bio-fuel resource. Over 90% of fuel ethanol is produced from maize (Wolt and Karaman, 2007). The United States Department of Energy (DOE) and United States Department of Agriculture (USDA) estimated that more than 30% of the U.S. transportation fuel needs could be displaced by renewable bio-fuel from cellulosic resources (Perlack et al., 2005). To reduce the U.S. reliance on foreign oil, the need for maize continues to increase. Using the maize seed as a production platform has an advantage that the maize grain waste after recombinant protein recovery can be used to produce fuel ethanol (Johnson et al., 2003). Typically, 99% of the grain mass remains as waste after the recombinant protein recovery (Paraman et al., 2010). Thus, the maize seed is an ideal production platform for coproduction of recombinant proteins and fuel ethanol in the United States. Strategies for expressing recombinant proteins in maize seeds are presented in Chapters 3, 4, 5, and 6 of this thesis. In addition to the advantages of using

the seed-targeting method described above, the maize stalk after harvest can be renewable bio-fuel resource.

1.1.6. Challenges

Although numerous studies have been published about producing recombinant proteins in maize seeds, there are still two major remaining challenges: 1) open-field release regulation of pharmaceutical maize and 2) low expression of recombinant products. Maize is a wind-pollinated plant, which means the pollen drifts with the wind and pollinates another plant. A major concern has been raised that transgenic maize could contaminate surrounding non-transgenic maize by pollen drift (Andow et al., 2004; Commandeur et al., 2003; Gressel, 2010). Open-field release of pharmaceutical plants is regulated by the U.S. Animal and Plant Health Inspection Service (APHIS). To overcome limitations of maize pollen drift, two strategies for producing pharmaceutical maize for open-field release are described in Chapter 5: 1) establishing biolistic- or *Agrobacterium*-mediated transformation methods of a tissue culture-amenable male-sterile line, and 2) conventional breeding strategies of introgressing of a transgene from male-fertile to male sterile transgenic maize..

The low expression of recombinant proteins in transgenic plants is the main challenge to plant molecular farming (Ma and Wang, 2012). To obtain high yield of plant-derived biogenic products, various approaches have been developed to enhance the expression or stability of recombinant proteins. These strategies include choosing suitable promoters, co-expressing with transcription factors, adding 5' and 3'UTR, optimizing codon usage, replacing signal peptide or other targeting sequences, adding ER

(endoplasmic reticulum) retention sequence, avoiding mRNA destabilizing sequences, fusing with endogenous proteins or tags, rebalancing seed storage protein, *etc.* (Egelkrout et al., 2012; Hood and Vicuna Requesens, 2011; Kermode, 2006; Kermode, 2012; Ma and Wang, 2012; Obembe et al., 2011). In Chapters 3, 4, and 6 of this thesis, strategies for improving recombinant protein expression in transgenic maize seeds are presented, including tissue-specific expression, promoter selection, protein fusion, ER retention, seed storage protein rebalancing, and co-expression of transcription factors.

1.1.7. Hypothesis and Objectives

The purpose of this dissertation research was to evaluate strategies that can lead to safe production of high yield and adequately modified recombinant proteins in maize seeds. We hypothesize that 1) maize has potential to generate mammalian-like exogenous proteins by co-expressing a transgene with an enzyme with post-translational modification activity; 2) small peptides, such as peptide biosurfactants, can be produced and purified with simple steps from maize seeds by fusing to tag proteins or maize endogenous proteins; 3) pharma corn (transgenic maize expressing proteins with pharmaceutical activities) can be safely and effectively produced in open fields by using male-sterile maize lines; and 4) recombinant protein expression enhanced in maize lines in which either production of one of the major seed protein is reduced or one of the transcription factors for seed protein is increased. To examine our hypotheses, the following four objectives were conducted:

1. Produce hydroxylated recombinant human collagen type I alpha 1 in maize seeds with co-expression of recombinant human prolyl 4-hydroxylase (Chapter 3)

2. Express the peptide biosurfactant GAM1 in maize seeds by fusing to the red fluorescent protein (RFP) or maize seed storage protein 22-kDa α-zein (Chapter 4)
3. Develop transformation of the tissue culture-amenable male-sterile maize and transfer the *E-coli* heat-labile enterotoxin B subunit (LT-B), a vaccine gene, from male-fertile to sterile-maize lines for open-field release (Chapter 5)
4. Evaluate the expression of green fluorescent protein (GFP) in maize lines in which either a 27 kDa γ-zein seed protein is silenced or a transcription factor prolamin box-binding factor (PBF) is overexpressed (Chapter 6)

1.2. Dissertation organization

This dissertation includes seven chapters. Chapter 1 is a general introduction. Chapter 2 is a literature review with a broad view of recombinant protein production in plants. The review includes: 1) a brief history of using plants to produce foreign proteins; 2) strategies to enhance the expression of recombinant proteins in plants; and 3) examples of producing exogenous proteins in transgenic maize. Chapters 1and 2 focus and explain the benefits and challenges of using transgenic plants, what kinds of plants are suitable, advantages of selecting tissue-specific expression and strategies to enhance production.

Chapters 3 through 6 are four independent, but related, studies in journal paper format. Chapter 3 presents the expression and characterization of recombinant human collagen type I α 1 (rCIα1) in transgenic maize. This work shows that rCIα1 not only can be produced in maize seed, but also can be post-translationally modified by co-expressing with a recombinant human prolyl 4-hydroxylase (rP4H). By taking advantage of high-

resolution mass spectrometry (HRMS), the rCIα1was determined a high proline hydroxylation ratio. The present study indicates that maize is able to produce recombinant protein with mammalian-like post-translational modifications. This work has been published in the journal ***BMC Biotechnology***. The gene construction work and maize transformation were completed by FibroGen Inc. and ProdiGene, Inc. We performed the protein and proteomics analyses on transgenic seeds that were transferred to Iowa State University, especially the hydroxylation analysis by using liquid chromatography tandem mass spectrometry (LC-MS/MS) on a Linear Ion Trap Orbitrap (LTQ Orbitrap) Mass Spectrometer. My major professor, Dr. Kan Wang, and Dr. Julio Baez from FibroGen. Inc, co-designed the experiments.

Chapter 4 describes efforts to express the biosurfactant peptide GMA1 in maize. Transgenic plants were successfully generated. However, the GAM1 was unable to be detected, possibly due to the low expression.

Chapter 5 is a paper about strategies to produce maize-derived pharmaceutical proteins by using male-sterile lines. This paper provides two strategies suitable for producing pharmaceutical proteins and avoiding cross-contamination by transgenic maize in open-field. This work has been published in the journal, ***Maydica***. Dr. Kan Wang designed and guided the experiments. My contributions to this paper were designing the breeding strategies, breeding and selecting the desired maize line in the greenhouse, analyzing the pollen viability, and conducting most field-release work.

Chapter 6 presents strategies to improve recombinant protein expression in maize by knocking down one major maize seed storage protein 27-kDa γ-zein and over expressing the maize seed protein transcription factor prolamin box-binding factor (PBF).

While some preliminary work is completed, this work is still in progress and will be submitted to a peer reviewed journal.

Chapter 7 is a general conclusion, summarizing all work described above. All projects presented in this dissertation are under the instruction and supervision by my major professor, Dr. Kan Wang. My POS committee members, Drs. Charles E Glatz, Martha James, Lawrence Johnson, and Paul Scott not only provided suggestions, but also shared materials and methods. Staff of the Protein Facility and the W.M. Keck Metabolomics Research Laboratory at Iowa State University helped me with the protein analysis performed for Chapters 3, 4 and 6. Dr. Lorena Moeller, who was graduate student and postdoctoral research associate in the Plant Transformation Facility, provided green fluorescent protein (GFP) transgenic corn and GFP enzyme-linked immunosorbent assay (ELISA) protocols used in Chapter 6. Bronwyn Frame, Marcy Main, Katey Warnberg and Rose Schick of the Plant Transformation Facility at Iowa State University did the maize transformation and greenhouse care for the constructs described in Chapters 4,5, and 6.

References

Andow D, Daniell H, Gepts P, Lamkey K, Nafziger E, Strayer D, Mellon M and Rissler J (2004) A growing concern: Protecting the food supply in an era of pharmaceutical and industrial crops: USC Publictions.

Barta, A., Sommergruber, K., Thompson, D., Hartmuth, K., Matzke, M. and Matzke, A. (1986) The expression of a nopaline synthase - human growth-hormone chimeric

gene in transformed tobacco and sunflower callus-tissue. *Plant Molecular Biology* **6**, 347-357.

Bevan, M., Flavell, R. and Chilton, M. (1983) A chimaeric antibiotic-resistance gene as a selectable marker for plant-cell transformation. *Nature* **304**, 184-187.

Boothe, J., Nykiforuk, C., Shen, Y., Zaplachinski, S., Szarka, S., Kuhlman, P., Murray, E., Morck, D. and Moloney, M. (2010) Seed-based expression systems for plant molecular farming. *Plant Biotechnology Journal* **8**, 588-606.

Chikwamba, R., Cunnick, J., Hathaway, D., McMurray, J., Mason, H. and Wang, K. (2002a) A functional antigen in a practical crop: LT-B producing maize protects mice against Escherichia coli heat labile enterotoxin (LT) and cholera toxin (CT). *Transgenic Research* **11**, 479-493.

Chikwamba, R., McMurray, J., Shou, H., Frame, B., Pegg, S., Scott, P., Mason, H. and Wang, K. (2002b) Expression of a synthetic E-coli heat-labile enterotoxin B sub-unit (LT-B) in maize. *Molecular Breeding* **10**, 253-265.

Clough, R., Pappu, K., Thompson, K., Beifuss, K., Lane, J., Delaney, D., Harkey, R., Drees, C., Howard, J. and Hood, E. (2006) Manganese peroxidase from the white-rot fungus Phanerochaete chrysosporium is enzymatically active and accumulates to high levels in transgenic maize seed. *Plant Biotechnology Journal* **4**, 53-62.

Cohen, S., Chang, A., Boyer, H. and Helling, R. (1973) Construction of biologically functional bacterial plasmids in-vitro. *Proceedings of the National Academy of Sciences of the United States of America* **70**, 3240-3244.

Commandeur, U., Twyman, R.M. and Fischer, R. (2003) The biosafety of molecular farming in plants. *AgBiotechNet* **5**, 110.

Conley, A., Joensuu, J., Jevnikar, A., Menassa, R. and Brandle, J. (2009) Optimization of Elastin-Like Polypeptide Fusions for Expression and Purification of Recombinant Proteins in Plants. *Biotechnology and Bioengineering* **103**, 562-573.

Egelkrout, E., Rajan, V. and Howard, J. (2012) Overproduction of recombinant proteins in plants. *Plant Science* **184**, 83-101.

Faye, L. and Gomord, V. (2010) Success stories in molecular farming-a brief overview. (vol 8, pg 525, 2010). *Plant Biotechnology Journal* **8**, 748-748.

Fischer, R. and Emans, N. (2000) Molecular farming of pharmaceutical proteins. *Transgenic Research* **9**, 279-299.

Gewin, V. (2003) Genetically modified corn - Environmental benefits and risks. *Plos Biology* **1**, 15-19.

Giddings, G. (2001) Transgenic plants as protein factories. *Current Opinion in Biotechnology* **12**, 450-454.

Gressel, J. (2010) Gene flow of transgenic seed-expressed traits: Biosafety considerations. *Plant Science* **179**, 630-634.

Hiatt, A., Cafferkey, R. and Bowdish, K. (1989) Production of antibodies in transgenic plants. *Nature* **342**, 76-78.

Hood, E., Bailey, M., Beifuss, K., Magallanes-Lundback, M., Horn, M., Callaway, E., Drees, C., Delaney, D., Clough, R. and Howard, J. (2003) Criteria for high-level expression of a fungal laccase gene in transgenic maize. *Plant Biotechnology Journal* **1**, 129-140.

Hood, E., Devaiah, S., Fake, G., Egelkrout, E., Teoh, K., Requesens, D., Hayden, C., Hood, K., Pappu, K., Carroll, J. and Howard, J. (2012) Manipulating corn

germplasm to increase recombinant protein accumulation. *Plant Biotechnology Journal* **10**, 20-30.

Hood, E., Love, R., Lane, J., Bray, J., Clough, R., Pappu, K., Drees, C., Hood, K., Yoon, S., Ahmad, A. and Howard, J. (2007) Subcellular targeting is a key condition for high-level accumulation of cellulase protein in transgenic maize seed. *Plant Biotechnology Journal* **5**, 709-719.

Hood, E. and Vicuna Requesens, D. (2011) Recombinant protein production in plants: challenges and solutions. *Methods in Molecular Biology: Recombinant Gene Expression Humana Press, New York*.

Hood, E., Witcher, D., Maddock, S., Meyer, T., Baszczynski, C., Bailey, M., Flynn, P., Register, J., Marshall, L., Bond, D., Kulisek, E., Kusnadi, A., Evangelista, R., Nikolov, Z., Wooge, C., Mehigh, R., Hernan, R., Kappel, W., Ritland, D., Li, C. and Howard, J. (1997) Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. *Molecular Breeding* **3**, 291-306.

Horvath, H., Huang, J., Wong, O., Kohl, E., Okita, T., Kannangara, C. and von Wettstein, D. (2000) The production of recombinant proteins in transgenic barley grains. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 1914-1919.

Johnson, L.A., May, J.B. and White, P.J. (2003) Wet milling: The basis for corn biorefineries. *Corn: Chemistry and Technology*, 449-494.

Kermode, A. (2006) Plants as factories for production of biopharmaceutical and bioindustrial proteins: lessons from cell biology. *Canadian Journal of Botany-Revue Canadienne De Botanique* **84**, 679-694.

Kermode, A.R. (2012) Seed Expression Systems for Molecular Farming. *Molecular Farming in Plants: Recent Advances and Future Prospects*, 89-123.

Kusnadi, A., Evangelista, R., Hood, E., Howard, J. and Nikolov, Z. (1998) Processing of transgenic corn seed and its effect on the recovery of recombinant beta-glucuronidase. *Biotechnology and Bioengineering* **60**, 44-52.

Lai, H., Engle, M., Fuchs, A., Keller, T., Johnson, S., Gorlatov, S., Diamond, M. and Chen, Q. (2010) Monoclonal antibody produced in plants efficiently treats West Nile virus infection in mice. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 2419-2424.

Lau, O. and Sun, S. (2009) Plant seeds as bioreactors for recombinant protein production. *Biotechnology Advances* **27**, 1015-1022.

Liu, C., Towler, M., Medrano, G., Cramer, C. and Weathers, P. (2009) Production of mouse interleukin-12 is greater in tobacco hairy roots grown in a mist reactor than in an airlift reactor. *Biotechnology and Bioengineering* **102**, 1074-1086.

Ma, J., Drake, P. and Christou, P. (2003) The production of recombinant pharmaceutical proteins in plants. *Nature Reviews Genetics* **4**, 794-805.

Ma, J., Hikmat, B., Wycoff, K., Vine, N., Chargelegue, D., Yu, L., Hein, M. and Lehner, T. (1998) Characterization of a recombinant plant monoclonal secretory antibody and preventive immunotherapy in humans. *Nature Medicine* **4**, 601-606.

Ma, S. and Wang, A. (2012) Molecular farming in plants: an overview. *Molecular Farming in Plants: Recent Advances and Future Prospects*, 1-20.

Mason, H., Lam, D. and Arntzen, C. (1992) Expression of hepatitis-b surface-antigen in transgenic plants. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 11745-11749.

Menassa, R., Hong, Z., Karatzas, C., Lazaris, A., Richman, A. and Brandle, J. (2004) Spider dragline silk proteins in transgenic tobacco leaves: accumulation and field production. *Plant Biotechnology Journal* **2**, 431-438.

Murray, C., Sutherland, P., Phung, M., Lester, M., Marshall, R. and Christeller, J. (2002) Expression of biotin-binding proteins, avidin and streptavidin, in plant tissues using plant vacuolar targeting sequences. *Transgenic Research* **11**, 199-214.

Obembe, O., Popoola, J., Leelavathi, S. and Reddy, S. (2011) Advances in plant molecular farming. *Biotechnology Advances* **29**, 210-222.

Paraman, I., Moeller, L., Scott, M., Wang, K., Glatz, C. and Johnson, L. (2010) Utilizing protein-lean coproducts from corn containing recombinant pharmaceutical proteins for ethanol production. *Journal of Agricultural and Food Chemistry* **58**, 10419-10425.

Parsons, J., Wirth, S., Dominguez, M., Bravo-Almonacid, F., Giulietti, A.M. and Rodriguez Talou, J. (2010) Production of human epidermal growth factor (hEGF) by in vitro cultures of nicotiana tabacum: effect of tissue differentiation and sodium nitroprusside addition. *Int J Biotechnol Biochem* **6**, 131-138.

Pen, J., Molendijk, L., Quax, W., Sijmons, P., Vanooyen, A., Vandenelzen, P., Rietveld, K. and Hoekema, A. (1992) Production of active bacillus-licheniformis alpha-

amylase in tobacco and its application in starch liquefaction. *Bio-Technology* **10**, 292-296.

Perlack, R.D., Wright, L.L., Turhollow, A.F., Graham, R.L., Stokes, B.J. and Erbach, D.C. (2005) Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply. DTIC Document.

Ramessar, K., Capell, T. and Christou, P. (2008) Molecular pharming in cereal crops. *Phytochemistry Reviews* **7**, 579-592.

Ruggiero, F., Exposito, J., Bournat, P., Gruber, V., Perret, S., Comte, J., Olagnier, B., Garrone, R. and Theisen, M. (2000) Triple helix assembly and processing of human collagen produced in transgenic tobacco plants. *Febs Letters* **469**, 132-136.

Scheller, J., Guhrs, K., Grosse, F. and Conrad, U. (2001) Production of spider silk proteins in tobacco and potato. *Nature Biotechnology* **19**, 573-577.

Shoji, Y., Chichester, J., Bi, H., Musiychuk, K., de la Rosa, P., Goldschmidt, L., Horsey, A., Ugulava, N., Palmer, G., Mett, V. and Yusibov, V. (2008) Plant-expressed HA as a seasonal influenza vaccine candidate. *Vaccine* **26**, 2930-2934.

Shoji, Y., Chichester, J., Jones, M., Manceva, S., Damon, E., Mett, V., Musiychuk, K., Bi, H., Farrance, C., Shamloul, M., Kushnir, N., Sharma, S. and Yusibov, V. (2011) Plant-based rapid production of recombinant subunit hemagglutinin vaccines targeting H1N1 and H5N1 influenza. *Human Vaccines* **7**, 41-50.

Sijmons, P., Dekker, B., Schrammeijer, B., Verwoerd, T., Vandenelzen, P. and Hoekema, A. (1990) Production of correctly processed human serum-albumin in transgenic plants. *Bio-Technology* **8**, 217-221.

Smith, H. and Wilcox, K. (1970) A restriction enzyme from hemophilus-influenzae .1. Purification and general properties. *Journal of Molecular Biology* **51**, 379-&.

Staub, J., Garcia, B., Graves, J., Hajdukiewicz, P., Hunter, P., Nehra, N., Paradkar, V., Schlittler, M., Carroll, J., Spatola, L., Ward, D., Ye, G. and Russell, D. (2000) High-yield production of a human therapeutic protein in tobacco chloroplasts. *Nature Biotechnology* **18**, 333-338.

Streatfield, S., Jilka, J., Hood, E., Turner, D., Bailey, M., Mayor, J., Woodard, S., Beifuss, K., Horn, M., Delaney, D., Tizard, I. and Howard, J. (2001) Plant-based vaccines: unique advantages. *Vaccine* **19**, 2742-2748.

Takaiwa, F. (2009) Allergen-specific immunotherapy with plant-based oral vaccines. *Immunotherapy* **1**, 517-519.

Tremblay, R., Wang, D., Jevnikar, A. and Ma, S. (2010) Tobacco, a highly efficient green bioreactor for production of therapeutic proteins. *Biotechnology Advances* **28**, 214-221.

Wilken, L. and Nikolov, Z. (2012) Recovery and purification of plant-made recombinant proteins. *Biotechnology Advances* **30**, 419-433.

Wolt, J. and Karaman, S. (2007) Estimated environmental loads of alpha-amylase from transgenic high-amylase maize. *Biomass & Bioenergy* **31**, 831-835.

Woodard, S., Mayor, J., Bailey, M., Barker, D., Love, R., Lane, J., Delaney, D., McComas-Wagner, J., Mallubhotla, H., Hood, E., Dangott, L., Tichy, S. and Howard, J. (2003) Maize (Zea mays)-derived bovine trypsin: characterization of the first large-scale, commercial protein product from transgenic plants. *Biotechnology and Applied Biochemistry* **38**, 123-130.

Xie, T., Qiu, Q., Zhang, W., Ning, T., Yang, W., Zheng, C., Wang, C., Zhu, Y. and Yang, D. (2008) A biologically active rhIGF-1 fusion accumulated in transgenic rice seeds can reduce blood glucose in diabetic mice via oral delivery. *Peptides* **29**, 1862-1870.

Xu, J., Dolan, M.C., Medrano, G., Cramer, C.L. and Weathers, P.J. (2011a) Green Factory: Plants as bioproduction platforms for recombinant proteins. *Biotechnology Advances*. **30**:1171-1184

Xu, J., Shpak, E., Gu, T., Moo-Young, M. and Kieliszewski, M. (2005) Production of recombinant plant gum with tobacco cell culture in bioreactor and gum characterization. *Biotechnology and Bioengineering* **90**, 578-588.

Xu, X., Gan, Q., Clough, R., Pappu, K., Howard, J., Baez, J. and Wang, K. (2011b) Hydroxylation of recombinant human collagen type I alpha 1 in transgenic maize co-expressed with a recombinant human prolyl 4-hydroxylase. *Bmc Biotechnology* **11**.

Xue, G., Patel, M., Johnson, J., Smyth, D. and Vickers, C. (2003) Selectable marker-free transgenic barley producing a high level of cellulase (1,4-beta-glucanase) in developing grains. *Plant Cell Reports* **21**, 1088-1094.

Yang, J., Barr, L., Fahnestock, S. and Liu, Z. (2005) High yield recombinant silk-like protein production in transgenic plants through protein targeting. *Transgenic Research* **14**, 313-324.

Yang, L., Tada, Y., Yamamoto, M., Zhao, H., Yoshikawa, M. and Takaiwa, F. (2006) A transgenic rice seed accumulating an anti-hypertensive peptide reduces the blood pressure of spontaneously hypertensive rats. *Febs Letters* **580**, 3315-3320.

Yang, L., Wakasa, Y. and Takaiwa, F. (2008) Biopharming to increase bioactive peptides in rice seed. *Journal of Aoac International* **91**, 957-964.

Zhang, C., Baez, J., Pappu, K. and Glatz, C. (2009a) Purification and characterization of a transgenic corn grain-derived recombinant collagen type I alpha 1. *Biotechnology Progress* **25**, 1660-1668.

Zhang, C., Glatz, C., Fox, S. and Johnson, L. (2009b) Fractionation of transgenic corn seed by dry and wet milling to recover recombinant collagen-related proteins. *Biotechnology Progress* **25**, 1396-1401.

CHAPTER 2: LITERATURE REVIEW

2.1. Approaches for recombinant protein expression in plants

2.1.1. Production platforms: Leaf- and seed-based systems

Plants have been used as important sources for producing recombinant proteins due to the development of plant transformation technology. Various proteins have been produced in both public and private laboratories during the past decades, including therapeutics, nutraceuticals, vaccine antigens, antibodies, and polymers (Arntzen et al., 2005; De Muynck et al., 2010; Giddings et al., 2000; Ling et al., 2010; Mooney, 2009; Sala et al., 2003; Sharma and Sharma, 2009; Stoger et al., 2005; Wilken and Nikolov, 2012). Plant production platforms for foreign proteins have been mainly agricultural crops such as tobacco, rice, and maize (Basaran and Rodrıguez-Cerezo, 2008; Fox, 2006; Wilken and Nikolov, 2012). Tobacco is the leading leaf-based protein expression system for commercial products (Twyman et al., 2003; Wilken and Nikolov, 2012). The first tobacco-derived commercial product was immunoglobulin A (IgA) (Ma et al., 1998). As a leaf-based system, tobacco can achieve high biomass yield with less pollen contamination risk than other seed-based systems, since the plant can be harvested before flowering (Wilken and Nikolov, 2012). Also, unlike the seed-based system, the process of extracting recombinant proteins from leafy plants does not require additional operations, such as seed grinding, soaking, and de-oiling (Schillberg et al., 2005). However, major disadvantages are significant in leafy plants like tobacco—high water content and low storage stability (De Muynck et al., 2010; Doran, 2006; Wilken and Nikolov, 2012). Although leaf-based systems are able to produce a high yield of

recombinant proteins, this yield may be offset by an increased instability of expressed proteins in metabolically active tissues, such as leaves (De Muynck et al., 2010; Doran, 2006).

The seed-based recombinant production platform offers several advantages over the leaf-based platform. Seeds are naturally designed for protein storage for long periods of time without or with very little degradation (Fischer et al., 2009; Kusnadi et al., 1997; Nikolov and Hammes, 2002). Seeds can be stored after harvest without cooling or immediate isolation that are required for leafy tissues (Boothe et al., 2010; Fiedler and Conrad, 1995; Hood et al., 1997; Lorenz and Kulp, 1991). The seed-based recombinant production platform has been used to produce various pharmaceutical or industrial products, including cytokines, antibodies, vaccines, and enzymes (Chikwamba et al., 2002a; Chikwamba et al., 2002b; Nochi et al., 2007; Stoger et al., 2000; Streatfield and Howard, 2003; Wu et al., 2007; Zhu et al., 1994). Compared to the leaf-based platform, seeds have a relatively lower biomass and higher cross-contamination risk by pollen drift for non-self-pollinated plants. However, considering the stability of foreign proteins, post-harvest processing, and overall cost, the seed-based platform is still more suitable for many recombinant proteins production on a large scale (Nikolov and Hammes, 2002; Schillberg et al., 2005).

2.2.1. Cis-elements considerations

2.2.1.1. Promoters

Gene expression in plants is regulated at different levels. The selection of the promoter is the first thing to consider before making the construct. Promoter elements can

dramatically affect the level of messenger RNA and, thus, influence the accumulation of the protein (Ma and Wang, 2012). To obtain high transcription levels, many constitutive promoters have been used to drive the transgene. The most widely used promoter is the cauliflower mosaic virus (CaMV) 35S promoter (Odell et al., 1985). This viral promoter is strong in many dicotyledonous plant species and has been used to produce a number of pharmaceutical proteins with relatively high expression (Tremblay et al., 2010; Twyman et al., 2003). However, the 35S promoter is much weaker in monocotyledonous plants, although it is still often used to drive maker genes for selection (Hood and Vicuna Requesens, 2011; Ma and Wang, 2012). In dicotyledonous plants, strong constitutive promoters, such as rice actin-1 or maize ubiquitin-1, are more frequently used for expressing foreign proteins (Streatfield, 2007; Twyman et al., 2003). To avoid toxicity to host plants from the exogenous genes or to obtain expression in a particular tissue of the plant avoiding protein degradation, tissue-specific promoters should be considered for use. Examples of tissue-specific promoters include zein (from maize), glutenin (from rice and wheat), and legumin (from pea) promoters (Ma et al., 2003). Many pharmaceutical proteins have been expressed with tissue- or organ-specific promoters, such as vaccine antigen HBsAg M, murine single chain variable fragment G4, human interferon-α, and the heat labile enterotoxin (LT) B subunit of *Escherichia coli* (*E. coli*) (Chikwamba et al., 2002a; Chikwamba et al., 2002b; De Jaeger et al., 2002; He et al., 2008). However, even the same promoter can result in different protein accumulation patterns within different species (Stoger et al., 2002). So, it is important to choose a suitable promoter for the specific system.

2.2.1.2. Codon usage

Different species have different codon usage ratios for a specific amino acid, according to the supply of the tRNAs for that amino acid (Ikemura, 1985). Codon usage should be considered when a transgene is expressed in a host plant with different codon use frequency to achieve optimal expression levels. Nowadays, codon optimization is a standard practice for transgene expression. Many codon optimization programs are available on the Internet based on the analysis of sequences available for a given species (Liu and Xue, 2005; Streatfield et al., 2001b): 1) DNAWorks http://helixweb.nih.gov/dnaworks/ (Hoover and Lubkowski, 2002); 2) Gene Designer https://www.dna20.com/ (Villalobos et al., 2006), 3) GeneDesign http://genedesign.thruhere.net/gd (Richardson et al., 2006), and 4) OPTIMIZER http://genomes.urv.cat/OPTIMIZER/ (Puigbo et al., 2007).

2.2.1.3. Subcellular targeting

Subcellular targeting is an alternative strategy to enhance the expression of exogenous proteins by avoiding protein degradation caused by proteases in the host plant. Recombinant proteins can be targeted to the cell wall, the vacuole, the mitochondria, and the chloroplast by fusing with different targeting signal. If a targeting signal sequence is not designed, the recombinant protein may be released in the cytoplasm where the protein could be degraded by proteases (Hood and Vicuna Requesens, 2011). The endoplasmic reticulum (ER) is the entry point for the protein secretory pathway, where many endogenous proteins or even foreign proteins can assemble (Lau and Sun, 2009; Nuttall et al., 2002). The ER is also a protective environment for recombinant proteins, since fewer proteases are found there (Hood and Vicuna Requesens, 2011). Additionally,

recombinant proteins can be modified at the post-translational level such as glycosylation (Ma et al., 2003). All these benefits make the ER an ideal target compartment for recombinant proteins. The ER retention sequences, such as KDEL (Lys-Asp-Glu-Leu) or HDEL (His-Asp-Glu-Leu), are used to keep recombinant proteins within the ER. This strategy has significantly enhanced the expression of foreign proteins in transgenic plants (Ko et al., 2003; Ma et al., 2005; Vaquero et al., 2002).

2.2.1.4. UTRs

The 5'- and 3'-UTRs (untranslated region) are important to the stability of the mRNA. Replacing the 5'-UTR and signal peptide of the human α-iduronidase gene with the arcelin gene increases mRNA and protein levels of α-iduronidase, and results in increased α-iduronidase's activity as well (Kermode, 2006). The extremely short 5'-UTRs (< 20 nucleotides) may inhibit the formation of the ribosome complex or the recognition of the start codon (Kawaguchi and Bailey-Serres, 2005; Kozak, 1991). The GC content is another crucial factor of the 5'-UTR. A low GC content (< 33 %) of the 5'-UTR is preferred while high GC content (> 50 %) inhibits ribosome loading (Kawaguchi and Bailey-Serres, 2005). Using the arcelin gene's 3'UTR and flanking sequences increase the activity of the α-iduronidase, as well as adding the 5'-UTR (Kermode, 2006). The highest expression level of the α-iduronidase was achieved by using the 5'- and 3'-UTRs, as well as the signal peptide sequence of the arcelin gene (Kermode, 2006). Although translation depends on the interaction between the 5' cap and 3' poly A, the 3'-UTR region is very critical to controling the stability of the mRNA for some plant genes (Chan and Yu, 1998; Gil and Green, 1996).

2.2.1.5. Protein fusion and tags

Protein fusion approaches are effective for boosting the expression, increasing solubility and stability, and facilitating isolation and purification of recombinant proteins (Hondred et al., 1999). Several fusion proteins have been developed and used to increase the production of recombinant proteins in plants. Zera is maize seed storage protein γ-zein derived proline-rich N-terminus domain, which has self-assembling and protein body formation properties (Ma and Wang, 2012). The expression of F1-V hybrid vaccine antigens increased 3 times, when fused with Zera in *N. benthamiana*, *Medicago sativa* (alfalfa) and *N. tabacum* NT1 cells (Alvarez et al., 2010). Another successful fusion partner example is elastin-like polypeptides (ELPs), which are artificially-designed biopolymers (Urry, 1992). ELP fusion proteins increase the expression of recombinant proteins in plants, including human IL-10 and murine interleukin-4, the full-size anti-human immunodeficiency virus type 1 antibody 2F5, and anti-foot and mouth disease virus single variable antibody fragments (Floss et al., 2008; Joensuu et al., 2009; Patel et al., 2007).

Affinity tags are used to facilitate the purification and quantification of recombinant proteins, especially when there is no antibody for the target protein. An ideal affinity tag should be small to allow the high yield and purity of the target protein after purification (Terpe, 2003). A widely used affinity tag example is the poly-histidine tag or 6XHis, since it usually consists of six consecutive histidines. The His-tagged proteins have inexpensively achieved high yield with large volume in bacterial systems, but are not very successful in plant systems due to the high background (Sharma and Sharma, 2009). Another frequently used affinity tag example is StrepII, consisting of eight amino

acids (WSHPQFEK). It was used to purify membrane-anchored protein kinase, NtCDPK2 from *N. benthamiana* with high yield by one-step using StrepII-Tactin column (Witte et al., 2004). Although affinity tags are usually very small, they still could affect the target protein's folding and normal functions. The FLAG tag system was designed to solve this problem. FLAG tag contains eight amino acids (DYKDDDDK) and can be released from the fusion protein by enterokinase treatment (Maroux et al., 1971). The purity of isolated FLAG-tagged protein was about 90% (Schuster et al., 2000). Other commonly used affinity tags include c-myc, glutathione S-transferase, calmodulin-binding peptide, maltose-binding protein, and the cellulose-binding domain (Terpe, 2003). There are numerous purification and quantification kits commercially available for these tags, which facilitate the expression and down-stream characterization of recombinant proteins.

2.2.2. Transcription factors

2.2.2.1. Regulation in plants by transcription factors

Plant genes are under the control of different factors like most eukaryotes. Transcription factors, known as a group of proteins recognizing particular DNA sequences in the promoter region of genes, usually affect the transcription of genes through protein-protein interaction activating RNA polymerase II (Grotewold, 2008). Based on the conserved folds in the DNA binding domains, transcription factors are categorized into more than fifty different families (Riechmann and Ratcliffe, 2000).

The regulatory activity of most transcription factors are affected by histone modifications and the chromatin structure (Kouzarides, 2007). With increasing

knowledge, it has been realized the regulation of the transcription factors are complicated when associated with the plant metabolism network, especially related to chromatin functions (Hernandez et al., 2007).

With the development of plant functional genomics, gene regulation by transcription factors has been discovered—especially in *Arabidopsis thaliana*, rice, and maize (Alonso and Ecker, 2006; Pastori and Foyer, 2002; Shinozaki and Yamaguchi-Shinozaki, 2000). In *Arabidopsis thaliana*, many transcription factors, as well as their regulating metabolic pathways, have been identified, for instance, PAP1/PAP2 (R2R3-MYB) regulate anthocyanin biosynthesis (Borevitz et al., 2000) and AtMYB12/AtMYB11/AtMYB111 regulate flavonol accumulation (Mehrtens et al., 2005; Stracke et al., 2007). TT2 (R2R3-MYB) and TT8 (bHLH) together are under the regulation of seed coat proanthocyanidins (Nesi et al., 2000; Nesi et al., 2001). In maize, it was determined P1 (R2R3-MYB) regulates phlobaphene accumulation (Grotewold et al., 1994), and C1/PL1 (R2R3-MYB) and R/B (bHLH) together are under the regulation of anthocyanins (Hernandez et al., 2004).

Since transcription factors are involved in many plant metabolic pathways, regulations, strategies, or technologies using transcription factor-based mechanisms will be part of the next generation of successful biotechnology-derived plants (Century et al., 2008). The production of isoflavones was increased in tobacco, *Arabidopsis thaliana*, and maize by over-expression of PAP1 and anthocyanidin reductase (ANR) (Xie et al., 2006), and co-expression of the flavonoid regulators with isoflavone synthase (IFS) (Yu et al., 2000). The expression of isoflavones was also successfully boosted in soybean by

expressing a synthetic regulator CRC (fused from maize C1 and R) and eliminating the endogenous F3H (Yu et al., 2003).

To enhance the production of recombinant proteins in transgenic maize, using transcription factors is one of the strategies presented in this dissertation. Two maize seed protein transcription factors, O2 and PBF, are described in detail below.

2.2.2.2. Transcription factors in maize: O2

In the 1920s, a natural maize mutant *opaque-2* (*o2*) was discovered with a seed phenotype change from glassy to opaque under light (Mertz et al., 1964). Major maize seed storage proteins, zeins, are reduced by 60 to 80% in *o2* (Aukerman et al., 1991). In maize seed endosperm, about 60% of the total protein are zeins (Mertz et al., 1964). Zeins can be classified into four groups—α-, β-, γ, and δ-zeins. Zein reduction in *o2* resulted from the suppression of the 22-kDa α-zein, so it was hypothesized the 22-kDa α-zein gene is regulated by O2, which may encode a transcription factor activating the 22-kDa α-zein gene expression (Burr and Burr, 1982; Kodrzycki et al., 1989; Pedersen et al., 1980; Soave et al., 1981).

The O2 gene was identified and cloned (Motto et al., 1988; Schmidt et al., 1987). The cDNA of O2 was sequenced, indicating the O2 encoded protein consists of a leucine-zipper DNA binding motif (Hartings et al., 1989; Landschulz et al., 1988; Schmidt et al., 1990), which recognizes the specific sequence of the target DNA (Neuberg et al., 1989; Turner and Tjian, 1989). This domain with the leucine repeat has been named bZip domain (Vinson et al., 1989). O2 is able to bind with the promoter of the 22-kDa α-zein gene and the bZip domain in O2 is responsible for this binding (Aukerman et al., 1991; Schmidt et al., 1990). The O2 binding site in the 22-kDa α-zein DNA gene was identified

and only found in the 22-kDa α-zein gene and not other distantly related zein genes, which explains the specific effect on the expression of 22-kDa α-zein in the *o2* mutant (Schmidt et al., 1992). Non-zein proteins increased in *o2*, which improved the nutrition of maize seeds, since zeins lack lysine, an essential amino acid (Coleman and Larkins, 1999; Osborne et al., 1914). However, the *o2* mutant is more susceptible to diseases associated with low yield, which limit further applications.

2.2.2.3. Transcription factors in maize: PBF

Prolamins are the most abundant cereal storage proteins with high proline content and found in many grains, such as wheat (gliadin), barley (hordein), rye (secalin), sorghum (kafirin), and maize (zein) (Shewry and Halford, 2002). Prolamins are generally soluble only in strong alcohol solutions. The expression of prolamins appears to be highly regulated by both cis-elements and transcription factors (Marzabal et al., 2008). Several endosperm boxes are conserved in prolamin genes. In maize, a 7 bp prolamin box has been identified: p-box (5'-TG(T/C/A)AAAG-3') (Boronat et al., 1986; Ottoboni et al., 1993; Thompson and Larkins, 1989). The P-box motif is recognized by the Dof transcription factors so-called PBFs (prolamin-box binding factors) (Mena et al., 1998; VicenteCarbajosa et al., 1997; Yanagisawa, 2004).

Dof (DNA binding with one finger) domain proteins are transcription factors specifically found in plants with highly conserved DNA-binding domains, including a single C_2-C_2 zinc finger (Umemura et al., 2004; Yanagisawa, 2002; Yanagisawa, 2004). The maize PBF interacts with prolamin box motifs in vitro with maize 22-kDa α-zein gene promoter (VicenteCarbajosa et al., 1997). The maize 27-kDa γ-zein promoter contains four prolamin boxes (Pb1, Pb2, Pb3, and Pb4) (Marzabal et al., 1998), which

could possibly be recognized by PBF. A recent study confirmed that only the 27-kDa γ- and 22-kDa α-zein gene expressions were affected when the PBF expression was knocked down by RNAi (RNA interference) (Wu and Messing, 2012a). By co-expression with PBF, the 27-kDa γ-zein promoter has been enhanced seven times stronger in tobacco (Marzabal et al., 2008). Thus, using transcription factors to enhance the target promoter can be a good strategy to increase the expression of recombinant proteins in plants.

2.2.3. Major seed protein gene silencing

Plant seeds have the ability to rebalance nutrition components after some major protein is eliminated. In the maize *o2* mutant, non-zein proteins have increased, while zeins are reduced significantly, resulting in improved maize seed nutritional quality (Coleman and Larkins, 1999). In cotton, the RNAi mediated knock-down of two key fatty acid desaturase genes leads to the increase of high-oleic and high-stearic cotton seed oil, essential fatty acids for human heart health (Liu et al., 2002). In soybean, while the expression of α and α' subunits of β-conglycinin is suppressed by sequence-mediated gene silencing in seeds, another soybean seed storage protein glycinin is increased to balance the total seed protein (Kinney et al., 2001), and the glycinin driven GFP expression was increased by 3.5-4.0 fold when introgressed into the β-conglycinin suppression line compared to the parent line (Schmidt and Herman, 2008).

Therefore, a strategy is raised to improve the nutritional quality or to enhance the expression of recombinant proteins in seeds, using the gene silencing mechanism to reduce a major seed protein gene (Kermode, 2006). In *Arabidopsis thaliana*, a transgene *arcelin5-I* from *Phaseolus vulgaris* increased up to 15% of the total soluble protein of the

seed when the endogenous seed storage protein 2S albumin is silenced by antisense (Goossens et al., 1999). A similar approach was also achieved in rice suspension culture. The recombinant human granulocyte-macrophage colony-stimulating factor (rhGM-CSF) increased by 1.9 fold with 22% reduction of α-amylase by RNAi (Kim et al., 2008).

Two major maize seed storage proteins, α- and γ-zeins, have been silenced by antisense or RNAi methods (Huang et al., 2004; Segal et al., 2003; Wu et al., 2010; Wu and Messing, 2012b). Non-zein proteins increased in these transgenic maize lines and seed nutritional quality improved high-lysine maize (Houmard et al., 2007). However, although recombinant proteins are commonly produced in transgenic maize, the seed protein gene silencing strategy has yet to be been adopted to improve the expression of recombinant proteins in transgenic maize. Chapter 5 focuses on use of such a strategy.

2.2. Case study: Recombinant protein expression in transgenic maize

2.2.1. Human collagen type I α 1

Collagens are the most abundant proteins in mammals and have been widely used in food and pharmaceutical industries (Olsen et al., 2003). More than one-half million metric tons of collagen and its denatured form, gelatin, are used in medical applications annually (Ritala et al., 2008). Current sources of collagen and gelatin are all animal collagen-rich tissues (Eskelin et al., 2009). With increasing public concerns for animal-derived collagen and gelatin safety issues (Asher, 1999), alternative non-animal-derived recombinant collagen expression systems have been developed, such as mammalian cells (Geddis and Prockop, 1993), insect cell cultures (Myllyharju et al., 1997), yeast (Vuorela et al., 1997), and plant cell culture (Ritala et al., 2008).

With advantages of low cost, high capacity, low contamination risk, and inexpensive storage, the transgenic plant system would be suitable for producing recombinant collagens (Ma et al., 2003; Menkhaus et al., 2004). Human collagen type I is the most abundant collagen type in the human body, containing two α 1 (CIα1) and one α 2 (CIα2) chains (Van der Rest and Garrone, 1991). The recombinant human collagen type I α 1 (rCIα1) has been expressed in tobacco (Ruggiero et al., 2000), barley (Eskelin et al., 2009), and maize (Xu et al., 2011; Zhang et al., 2009a; Zhang et al., 2009b).

The mature human CIα1 in human bodies is hydroxylated by human prolyl 4-hydroxylase (Van der Rest and Garrone, 1991). Thus, it is a challenge to express rCIα1 in the non-mammalian expression system. Plants have no or very low hydroxylation activity on CIα1. In Chapter 2, we present work showing maize has the ability to hydroxylate rCIα1 if recombinant human prolyl 4-hydroxylase is co-expressed with collagen.

2.2.2. GAM1

Biosurfactants are bio-derived molecules that lower the surface tension between two liquids, or between a liquid and a solid (Das et al., 2008). Biosurfactants are sustainable alternatives to fossil fuel-based surfactants (Clapes and Infante, 2002; Kaar et al., 2009; Van Hamme et al., 2006). Petroleum-based surfactants are not environmentally friendly and their prices are directly affected by oil prices. Therefore, in recent decades biosurfactants have been developed and applied in biomedical, pharmaceutical, environmental, and food industries (Banat et al., 2010; Nitschke and Costa, 2007; Singh et al., 2007).

GAM1 is an interfacially active designed peptide biosurfactant. Such peptide biosurfactants have been recently reported to have advantages of lowering the interfacial tension and control of interfacial rheology (Dexter et al., 2006; Dexter and Middelberg, 2007; Malcolm et al., 2006). GAM1 is a 22 amino acid peptide (sequence GMKQLADSLHQLARQVSRLEHA) modified from a previously reported stimuli-responsive peptide biosurfactant AM1 (Dexter et al., 2006; Dexter and Middelberg, 2008; Kaar et al., 2009). GAM1 has been successfully expressed in *E. coli* as a fusion to a maltose-binding protein with surface activity and the ability to trigger a switch between form-stabilizing and form-destabilizing states (Kaar et al., 2009). In Chapter 3, attempts to produce GMA1 in transgenic maize via fusion of the red fluorescent protein (RFP) and maize seed endogenous protein 22-kDa α-zein are presented.

2.2.3. Heat-labile enterotoxin B subunit (LT-B)

Enterotoxigenic *E. coli* (ETEC) colonizes in the large intestine of human beings (Hudault et al., 2001) and can cause severe diarrhea (Williams et al., 1999). The heat-labile enterotoxin from *E. coli* is about 84 kDa containing a 27 kDa toxic A subunit non-covalently linked to a ring shaped pentamer composed by five identical non-toxic 11 kDa B subunits. Thus, the B subunit of the heat-labile enterotoxin (LT-B) is an excellent choice as a potential vaccine against the toxin causing diarrhea, since the B subunit can induce secretory IgA and serum IgG (Millar et al., 2001). The recombinant LT-B has been produced in a variety of plants, including transgenic maize, tobacco, potato, and ginseng plants (Chikwamba et al., 2002a; Chikwamba et al., 2002b; Kang et al., 2006; Kang et al., 2003; Mason et al., 1998; Streatfield et al., 2001a). Especially, the benefits of

producing antigens as edible vaccines have been shown (Ramessar et al., 2008; Streatfield, 2007). The feeding experiments with mice indicated diarrhea symptoms induced by *E. coli* toxins in mice could be reduced by feeding maize-derived LT-B (Chikwamba et al., 2002a).

However, the open-field released transgenic plants producing pharmaceuticals, such as the LT-B, may meet challenges from potential health and environmental risks associated with its use. The major concern is whether maize-derived antigens will contaminate the food chain (Ripplinger et al., 2009). The pollen drift from maize could be considered a potential contamination source. Thus, in Chapter 4, we present strategies to overcome this challenge by using a tissue culture-amenable male-sterile maize line.

References

Alonso, J. and Ecker, J. (2006) Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in Arabidopsis. *Nature Reviews Genetics* **7**, 524-536.

Alvarez, M., Topal, E., Martin, F. and Cardineau, G. (2010) Higher accumulation of F1-V fusion recombinant protein in plants after induction of protein body formation. *Plant Molecular Biology* **72**, 75-89.

Arntzen, C., Plotkin, S. and Dodet, B. (2005) Preface - Plant-derived vaccines and antibodies: potential and limitations. *Vaccine* **23**, 1753-1756.

Asher, D.M. (1999) The transmissible spongiform encephalopathy agents: concerns and responses of United States regulatory agencies in maintaining the safety of biologics. *Developments in Biological Standardization* **100**, 103.

Aukerman, M.J., Schmidt, R.J., Burr, B. and Burr, F.A. (1991) An arginine to lysine substitution in the bZIP domain of an opaque-2 mutant in maize abolishes specific DNA binding. *Genes & Development* **5**, 310-320.

Banat, I., Franzetti, A., Gandolfi, I., Bestetti, G., Martinotti, M., Fracchia, L., Smyth, T. and Marchant, R. (2010) Microbial biosurfactants production, applications and future potential. *Applied Microbiology and Biotechnology* **87**, 427-444.

Basaran, P. and Rodrıguez-Cerezo, E. (2008) Plant molecular farming: opportunities and challenges. *Critical reviews in Biotechnology* **28**, 153-172.

Boothe, J., Nykiforuk, C., Shen, Y., Zaplachinski, S., Szarka, S., Kuhlman, P., Murray, E., Morck, D. and Moloney, M. (2010) Seed-based expression systems for plant molecular farming. *Plant Biotechnology Journal* **8**, 588-606.

Borevitz, J., Xia, Y., Blount, J., Dixon, R. and Lamb, C. (2000) Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* **12**, 2383-2393.

Boronat, A., Martinez, M.C., Reina, M., Puigdomenech, P. and Palau, J. (1986) Isolation and sequencing of a 28 kD glutelin-2 gene from maize. common elements in the 5 鈥?flanking regions among zein and glutelin genes. *Plant Science* **47**, 95-102.

Burr, F. and Burr, B. (1982) 3 mutations in zea-mays affecting zein accumulation - a comparison of zein polypeptides, invitro synthesis and processing, messenger-rna levels, and genomic organization. *Journal of Cell Biology* **94**, 201-206.

Century, K., Reuber, T. and Ratcliffe, O. (2008) Regulating the regulators: The future prospects for transcription-factor-based agricultural biotechnology products. *Plant Physiology* **147**, 20-29.

Chan, M. and Yu, S. (1998) The 3 ' untranslated region of a rice alpha-amylase gene mediates sugar-dependent abundance of mRNA. *Plant Journal* **15**, 685-695.

Chikwamba, R., Cunnick, J., Hathaway, D., McMurray, J., Mason, H. and Wang, K. (2002a) A functional antigen in a practical crop: LT-B producing maize protects mice against Escherichia coli heat labile enterotoxin (LT) and cholera toxin (CT). *Transgenic Research* **11**, 479-493.

Chikwamba, R., McMurray, J., Shou, H., Frame, B., Pegg, S., Scott, P., Mason, H. and Wang, K. (2002b) Expression of a synthetic E-coli heat-labile enterotoxin B sub-unit (LT-B) in maize. *Molecular Breeding* **10**, 253-265.

Clapes, P. and Infante, M. (2002) Amino acid-based surfactants: Enzymatic synthesis, properties and potential applications. *Biocatalysis and Biotransformation* **20**, 215-233.

Coleman, C.E. and Larkins, B.A. (1999) 6. The Prolamins of Maize. *seed Proteins*, 109.

Das, P., Mukherjee, S. and Sen, R. (2008) Antimicrobial potential of a lipopeptide biosurfactant derived from a marine Bacillus circulans. *Journal of Applied Microbiology* **104**, 1675-1684.

De Jaeger, G., Scheffer, S., Jacobs, A., Zambre, M., Zobell, O., Goossens, A., Depicker, A. and Angenon, G. (2002) Boosting heterologous protein production in transgenic dicotyledonous seeds using Phaseolus vulgaris regulatory sequences. *Nature Biotechnology* **20**, 1265-1268.

De Muynck, B., Navarre, C. and Boutry, M. (2010) Production of antibodies in plants: status after twenty years. *Plant Biotechnology Journal* **8**, 529-563.

Dexter, A., Malcolm, A. and Middelberg, A. (2006) Reversible active switching of the mechanical properties of a peptide film at a fluid-fluid interface. *Nature Materials* **5**, 502-506.

Dexter, A. and Middelberg, A. (2007) Switchable peptide surfactants with designed metal binding capacity. *Journal of Physical Chemistry C* **111**, 10484-10492.

Dexter, A. and Middelberg, A. (2008) Peptides as functional surfactants. *Industrial & Engineering Chemistry Research* **47**, 6391-6398.

Doran, P. (2006) Foreign protein degradation and instability in plants and plant tissue cultures. *Trends in Biotechnology* **24**, 426-432.

Eskelin, K., Ritala, A., Suntio, T., Blumer, S., Holkeri, H., Wahlstrom, E., Baez, J., Makinen, K. and Maria, N. (2009) Production of a recombinant full-length collagen type I alpha-1 and of a 45-kDa collagen type I alpha-1 fragment in barley seeds. *Plant Biotechnology Journal* **7**, 657-672.

Fiedler, U. and Conrad, U. (1995) High-level production and long-term storage of engineered antibodies in transgenic tobacco seeds. *Bio-Technology* **13**, 1090-1093.

Fischer, R., Schillberg, S. and Twyman, R.M. (2009) Molecular farming of antibodies in plants. *Recent Advances in Plant Biotechnology New York: Springer*, 35-63.

Floss, D., Sack, M., Stadlmann, J., Rademacher, T., Scheller, J., Stoger, E., Fischer, R. and Conrad, U. (2008) Biochemical and functional characterization of anti-HIV antibody-ELP fusion proteins from transgenic plants. *Plant Biotechnology Journal* **6**, 379-391.

Fox, J. (2006) Turning plants into protein factories. *Nature Biotechnology* **24**, 1191-1193.

Geddis, A. and Prockop, D. (1993) Expression of human col1a1 gene in stably transfected ht-1080 cells - the production of a thermostable homotrimer of type-1 collagen in a recombinant system. *Matrix* **13**, 399-405.

Giddings, G., Allison, G., Brooks, D. and Carter, A. (2000) Transgenic plants as factories for biopharmaceuticals. *Nature Biotechnology* **18**, 1151-1155.

Gil, P. and Green, P. (1996) Multiple regions of the Arabidopsis SAUR-AC1 gene control transcript abundance: The 3' untranslated region functions as an mRNA instability determinant. *Embo Journal* **15**, 1678-1686.

Goossens, A., Van Montagu, M. and Angenon, G. (1999) Co-introduction of an antisense gene for an endogenous seed storage protein can increase expression of a transgene in Arabidopsis thaliana seeds. *Febs Letters* **456**, 160-164.

Grotewold, E. (2008) Transcription factors for predictive plant metabolic engineering: are we there yet? *Current Opinion in Biotechnology* **19**, 138-144.

Grotewold, E., Drummond, B., Bowen, B. and Peterson, T. (1994) The myb-homologous p-gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. *Cell* **76**, 543-553.

Hartings, H., Maddaloni, M., Lazzaroni, N., Difonzo, N., Motto, M., Salamini, F. and Thompson, R. (1989) The o2 gene which regulates zein deposition in maize endosperm encodes a protein with structural homologies to transcriptional activators. *Embo Journal* **8**, 2795-2801.

He, Z., Jiang, X., Qi, Y. and Luo, D. (2008) Assessment of the utility of the tomato fruit-specific E8 promoter for driving vaccine antigen expression. *Genetica* **133**, 207-214.

Hernandez, J., Heine, G., Irani, N., Feller, A., Kim, M., Matulnik, T., Chandler, V. and Grotewold, E. (2004) Different mechanisms participate in the R-dependent activity of the R2R3 MYB transcription factor C1. *Journal of Biological Chemistry* **279**, 48205-48213.

Hernandez, J.M., Feller, A., Morohashi, K., Frame, K. and Grotewold, E. (2007) The basic helix–loop–helix domain of maize R links transcriptional regulation and histone modifications by recruitment of an EMSY-related factor. *Proceedings of the National Academy of Sciences* **104**, 17222-17227.

Hondred, D., Walker, J., Mathews, D. and Vierstra, R. (1999) Use of ubiquitin fusions to augment protein expression in transgenic plants. *Plant Physiology* **119**, 713-723.

Hood, E. and Vicuna Requesens, D. (2011) Recombinant protein production in plants: challenges and solutions. *Methods in Molecular Biology: Recombinant Gene Expression Humana Press, New York*.

Hood, E., Witcher, D., Maddock, S., Meyer, T., Baszczynski, C., Bailey, M., Flynn, P., Register, J., Marshall, L., Bond, D., Kulisek, E., Kusnadi, A., Evangelista, R., Nikolov, Z., Wooge, C., Mehigh, R., Hernan, R., Kappel, W., Ritland, D., Li, C. and Howard, J. (1997) Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. *Molecular Breeding* **3**, 291-306.

Hoover, D.M. and Lubkowski, J. (2002) DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Research* **30**, e43-e43.

Houmard, N.M., Mainville, J.L., Bonin, C.P., Huang, S., Luethy, M.H. and Malvar, T.M. (2007) High lysine corn generated by endosperm 欽恠 pecific suppression of lysine catabolism using RNAi. *Plant biotechnology journal* **5**, 605-614.

Huang, S., Adams, W., Zhou, Q., Malloy, K., Voyles, D., Anthony, J., Kriz, A. and Luethy, M. (2004) Improving nutritional quality of maize proteins by expressing sense and antisense zein genes. *Journal of Agricultural and Food Chemistry* **52**, 1958-1964.

Hudault, S., Guignot, J. and Servin, A.L. (2001) Escherichia coli strains colonising the gastrointestinal tract protect germfree mice againstSalmonella typhimuriuminfection. *Gut* **49**, 47-55.

Ikemura, T. (1985) Codon usage and tRNA content in unicellular and multicellular organisms. *Molecular Biology and Evolution* **2**, 13-34.

Joensuu, J., Brown, K., Conley, A., Clavijo, A., Menassa, R. and Brandle, J. (2009) Expression and purification of an anti-Foot-and-mouth disease virus single chain variable antibody fragment in tobacco plants. *Transgenic Research* **18**, 685-696.

Kaar, W., Hartmann, B., Fan, Y., Zeng, B., Lua, L., Dexter, A., Falconer, R. and Middelberg, A. (2009) Microbial Bio-Production of a Recombinant Stimuli-Responsive Biosurfactant. *Biotechnology and Bioengineering* **102**, 176-187.

Kang, T., Lee, W., Choi, E., Kim, J., Kim, B. and Yang, M. (2006) Mass production of somatic embryos expressing Escherichia coli heat-labile enterotoxin B subunit in Siberian ginseng. *Journal of Biotechnology* **121**, 124-133.

Kang, T., Loc, N., Jang, M., Jang, Y., Kim, Y., Seo, J. and Yang, M. (2003) Expression of the B subunit of E-coli heat-labile enterotoxin in the chloroplasts of plants and its characterization. *Transgenic Research* **12**, 683-691.

Kawaguchi, R. and Bailey-Serres, J. (2005) mRNA sequence features that contribute to translational regulation in Arabidopsis. *Nucleic Acids Research* **33**, 955-965.

Kermode, A. (2006) Plants as factories for production of biopharmaceutical and bioindustrial proteins: lessons from cell biology. *Canadian Journal of Botany-Revue Canadienne De Botanique* **84**, 679-694.

Kim, N., Kim, T., Jang, Y., Shin, Y., Kwon, T. and Yang, M. (2008) Amylase gene silencing by RNA interference improves recombinant hGM-CSF production in rice suspension culture. *Plant Molecular Biology* **68**, 369-377.

Kinney, A., Jung, R. and Herman, E. (2001) Cosuppression of the alpha subunits of beta-conglycinin in transgenic soybean seeds induces the formation of endoplasmic reticulum-derived protein bodies. *Plant Cell* **13**, 1165-1178.

Ko, K., Tekoah, Y., Rudd, P., Harvey, D., Dwek, R., Spitsin, S., Hanlon, C., Rupprecht, C., Dietzschold, B., Golovkin, M. and Koprowski, H. (2003) Function and glycosylation of plant-derived antiviral monoclonal antibody. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 8013-8018.

Kodrzycki, R., Boston, R.S. and Larkins, B.A. (1989) The opaque-2 mutation of maize differentially reduces zein gene transcription. *The Plant Cell Online* **1**, 105-114.

Kouzarides, T. (2007) Chromatin modifications and their function. *Cell* **128**, 693.

Kozak, M. (1991) A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. *Gene Expression* **1**, 111-115.

Kusnadi, A., Nikolov, Z. and Howard, J. (1997) Production of recombinant proteins in transgenic plants: Practical considerations. *Biotechnology and Bioengineering* **56**, 473-484.

Landschulz, W.H., Johnson, P.F. and McKnight, S.L. (1988) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* **240**, 1759-1764.

Lau, O. and Sun, S. (2009) Plant seeds as bioreactors for recombinant protein production. *Biotechnology Advances* **27**, 1015-1022.

Ling, H., Pelosi, A. and Walmsley, A. (2010) Current status of plant-made vaccines for veterinary purposes. *Expert Review of Vaccines* **9**, 971-982.

Liu, Q., Singh, S. and Green, A. (2002) High-oleic and high-stearic cottonseed oils: Nutritionally improved cooking oils developed using gene silencing. *Journal of the American College of Nutrition* **21**, 205S-211S.

Liu, Q. and Xue, Q. (2005) Comparative studies on codon usage pattern of chloroplasts and their host nuclear genes in four plant species. *Journal of genetics* **84**, 55-62.

Lorenz, K.J. and Kulp, K. (1991) *Handbook of Cereal Science And Technology*. New York:Marcel Dekker.

Ma, J., Drake, P. and Christou, P. (2003) The production of recombinant pharmaceutical proteins in plants. *Nature Reviews Genetics* **4**, 794-805.

Ma, J., Hikmat, B., Wycoff, K., Vine, N., Chargelegue, D., Yu, L., Hein, M. and Lehner, T. (1998) Characterization of a recombinant plant monoclonal secretory antibody and preventive immunotherapy in humans. *Nature Medicine* **4**, 601-606.

Ma, S., Huang, Y., Davis, A., Yin, Z., Mi, Q., Menassa, R., Brandle, J. and Jevnikar, A. (2005) Production of biologically active human interleukin-4 in transgenic tobacco and potato. *Plant Biotechnology Journal* **3**, 309-318.

Ma, S. and Wang, A. (2012) Molecular Farming in Plants: An Overview. *Molecular Farming in Plants: Recent Advances and Future Prospects*, 1-20.

Malcolm, A., Dexter, A. and Middelberg, A. (2006) Foaming properties of a peptide designed to form stimuli-responsive interfacial films. *Soft Matter* **2**, 1057-1066.

Maroux, S., Baratti, J. and Desnuelle, P. (1971) Purification and specificity of porcine enterokinase. *Journal of Biological Chemistry* **246**, 5031-5039.

Marzabal, P., Busk, P., Ludevid, M. and Torrent, M. (1998) The bifactorial endosperm box of gamma-zein gene: characterisation and function of the Pb3 and GZM cis-acting elements. *Plant Journal* **16**, 41-52.

Marzabal, P., Gas, E., Fontanet, P., Vicente-Carbajosa, J., Torrent, M. and Ludevid, M. (2008) The maize Dof protein PBF activates transcription of gamma-zein during maize seed development. *Plant Molecular Biology* **67**, 441-454.

Mason, H., Haq, T., Clements, J. and Arntzen, C. (1998) Edible vaccine protects mice against Escherichia coli heat-labile enterotoxin (LT): potatoes expressing a synthetic LT-B gene. *Vaccine* **16**, 1336-1343.

Mehrtens, F., Kranz, H., Bednarek, P. and Weisshaar, B. (2005) The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiology* **138**, 1083-1096.

Mena, M., Vicente-Carbajosa, J., Schmidt, R. and Carbonero, P. (1998) An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and

activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm. *Plant Journal* **16**, 53-62.

Menkhaus, T., Bai, Y., Zhang, C., Nikolov, Z. and Glatz, C. (2004) Considerations for the recovery of recombinant proteins from plants. *Biotechnology Progress* **20**, 1001-1014.

Mertz, E., Nelson, O. and Bates, L. (1964) Mutant gene that changes protein composition and increases lysine content of maize endosperm. *Science* **145**, 279-&.

Millar, D., Hirst, T. and Snider, D. (2001) Escherichia coli heat-labile enterotoxin B subunit is a more potent mucosal adjuvant than its closely related homologue, the B subunit of cholera toxin. *Infection and Immunity* **69**, 3476-3482.

Mooney, B. (2009) The second green revolution? Production of plant-based biodegradable plastics. *Biochemical Journal* **418**, 219-232.

Motto, M., Maddaloni, M., Ponziani, G., Brembilla, M., Marotta, R., Difonzo, N., Soave, C., Thompson, R. and Salamini, F. (1988) Molecular-cloning of the ortho-2-m5 allele of zea-mays using transposon marking. *Molecular & General Genetics* **212**, 488-494.

Myllyharju, J., Lamberg, A., Notbohm, H., Fietzek, P., Pihlajaniemi, T. and Kivirikko, K. (1997) Expression of wild-type and modified pro alpha chains of human type I procollagen in insect cells leads to the formation of stable [alpha 1(I)](2)alpha 2(I) collagen heterotrimers and [alpha 1(I)](3) homotrimers but not [alpha 2(I)](3) homotrimers. *Journal of Biological Chemistry* **272**, 21824-21830.

Nesi, N., Debeaujon, I., Jond, C., Pelletier, G., Caboche, M. and Lepiniec, L. (2000) The TT8 gene encodes a basic helix-loop-helix domain protein required for expression

of DFR and BAN genes in Arabidopsis siliques. *The Plant Cell Online* **12**, 1863-1878.

Nesi, N., Jond, C., Debeaujon, I., Caboche, M. and Lepiniec, L. (2001) The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *Plant Cell* **13**, 2099-2114.

Neuberg, M., Schuermann, M., Hunter, J.B. and Moeller, R. (1989) Two functionally different regions in Fos are required for the sequence-specific DNA interaction of the Fos/Jun protein complex. *Nature* **338**, 589-590

Nikolov, Z. and Hammes, D. (2002) Production of recombinant proteins from transgenic crops. *Plants as Factories for Protein Production Kluwer Academic Publishers, Dordrecht*, 159-174.

Nitschke, M. and Costa, S. (2007) Biosurfactants in food industry. *Trends in Food Science & Technology* **18**, 252-259.

Nochi, T., Takagi, H., Yuki, Y., Yang, L., Masumura, T., Mejima, M., Nakanishi, U., Matsumura, A., Uozumi, A., Hiroi, T., Morita, S., Tanaka, K., Takaiwa, F. and Kiyono, H. (2007) Rice-based mucosal vaccine as a global strategy for cold-chain- and needle-free vaccination. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 10986-10991.

Nuttall, J., Vine, N., Hadlington, J., Drake, P., Frigerio, L. and Ma, J. (2002) ER-resident chaperone interactions with recombinant antibodies in transgenic plants. *European Journal of Biochemistry* **269**, 6042-6051.

Odell, J., Nagy, F. and Chua, N. (1985) Identification of dna-sequences required for activity of the cauliflower mosaic virus-35s promoter. *Nature* **313**, 810-812.

Olsen, D., Yang, C., Bodo, M., Chang, R., Leigh, S., Baez, J., Carmichael, D., Perala, M., Hamalainen, E., Jarvinen, M. and Polarek, J. (2003) Recombinant collagen and gelatin for drug delivery. *Advanced Drug Delivery Reviews* **55**, 1547-1567.

Osborne, T., Mendel, L., Ferry, E. and Wakeman, A. (1914) Nutritive properties of proteins of the maize kernel. *Journal of Biological Chemistry* **18**, 1-16.

Ottoboni, L., Leite, A., Yunes, J., Targon, M. and Arruda, P. (1993) Sequence-analysis of 22 kda-like alpha-coixin genes and their comparison with homologous zein and kafirin genes reveals highly conserved protein-structure and regulatory elements. *Plant Molecular Biology* **21**, 765-778.

Pastori, G. and Foyer, C. (2002) Common components, networks, and pathways of cross-tolerance to stress. The central role of "redox" and abscisic acid-mediated controls. *Plant Physiology* **129**, 460-468.

Patel, J., Zhu, H., Menassa, R., Gyenis, L., Richman, A. and Brandle, J. (2007) Elastin-like polypeptide fusions enhance the accumulation of recombinant proteins in tobacco leaves. *Transgenic Research* **16**, 239-249.

Pedersen, K., Bloom, K., Anderson, J., Glover, D. and Larkins, B. (1980) Analysis of the complexity and frequency of zein genes in the maize genome. *Biochemistry* **19**, 1644-1650.

Puigbo, P., Guzman, E., Romeu, A. and Garcia-Vallve, S. (2007) OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Research* **35**, W126-W131.

Ramessar, K., Sabalza, M., Capell, T. and Christou, P. (2008) Maize plants: An ideal production platform for effective and safe molecular pharming. *Plant Science* **174**, 409-419.

Richardson, S.M., Wheelan, S.J., Yarrington, R.M. and Boeke, J.D. (2006) GeneDesign: rapid, automated design of multikilobase synthetic genes. *Genome Research* **16**, 550-556.

Riechmann, J.L. and Ratcliffe, O.J. (2000) A genomic perspective on plant transcription factors. *Current Opinion in Plant Biology* **3**, 423-434.

Ripplinger, D., Hayes, D., Goggi, A. and Lamkey, K. (2009) Insuring against losses from transgenic contamination: The case of pharmaceutical maize. *American Journal of Agricultural Economics* **91**, 322-334.

Ritala, A., Wahlstrom, E., Holkeri, H., Hafren, A., Makelainen, K., Baez, J., Makinen, K. and Nuutila, A. (2008) Production of a recombinant industrial protein using barley cell cultures. *Protein Expression and Purification* **59**, 274-281.

Ruggiero, F., Exposito, J., Bournat, P., Gruber, V., Perret, S., Comte, J., Olagnier, B., Garrone, R. and Theisen, M. (2000) Triple helix assembly and processing of human collagen produced in transgenic tobacco plants. *Febs Letters* **469**, 132-136.

Sala, F., Rigano, M., Barbante, A., Basso, B., Walmsley, A. and Castiglione, S. (2003) Vaccine antigen production in transgenic plants: strategies, gene constructs and perspectives. *Vaccine* **21**, 803-808.

Schillberg, S., Twyman, R. and Fischer, R. (2005) Opportunities for recombinant antigen and antibody expression in transgenic plants - technology assessment. *Vaccine* **23**, 1764-1769.

Schmidt, M. and Herman, E. (2008) Proteome rebalancing in soybean seeds can be exploited to enhance foreign protein accumulation. *Plant Biotechnology Journal* **6**, 832-842.

Schmidt, R., Burr, F. and Burr, B. (1987) Transposon tagging and molecular analysis of the maize regulatory locus opaque-2. *Science* **238**, 960-963.

Schmidt, R.J., Burr, F.A., Aukerman, M.J. and Burr, B. (1990) Maize regulatory gene opaque-2 encodes a protein with a" leucine-zipper" motif that binds to zein DNA. *Proceedings of the National Academy of Sciences* **87**, 46-50.

Schmidt, R.J., Ketudat, M., Aukerman, M.J. and Hoschek, G. (1992) Opaque-2 is a transcriptional activator that recognizes a specific target site in 22-kD zein genes. *The Plant Cell Online* **4**, 689-700.

Schuster, M., Wasserbauer, E., Einhauer, A., Ortner, C., Jungbauer, A., Hammerschmid, F. and Werner, G. (2000) Protein expression strategies for identification of novel target proteins. *Journal of Biomolecular Screening* **5**, 89-97.

Segal, G., Song, R. and Messing, J. (2003) A new opaque variant of maize by a single dominant RNA-interference-inducing transgene. *Genetics* **165**, 387-397.

Sharma, A. and Sharma, M. (2009) Plants as bioreactors: Recent developments and emerging opportunities. *Biotechnology Advances* **27**, 811-832.

Shewry, P. and Halford, N. (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. *Journal of Experimental Botany* **53**, 947-958.

Shinozaki, K. and Yamaguchi-Shinozaki, K. (2000) Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Current Opinion in Plant Biology* **3**, 217-223.

Singh, A., Van Hamme, J. and Ward, O. (2007) Surfactants in microbiology and biotechnology: Part 2. Application aspects. *Biotechnology Advances* **25**, 99-121.

Soave, C., Reggiani, R., Difonzo, N. and Salamini, F. (1981) Clustering of genes for 20 kd zein subunits in the short arm of maize chromosome-7. *Genetics* **97**, 363-377.

Stoger, E., Ma, J., Fischer, R. and Christou, P. (2005) Sowing the seeds of success: pharmaceutical proteins from plants. *Current Opinion in Biotechnology* **16**, 167-173.

Stoger, E., Sack, M., Perrin, Y., Vaquero, C., Torres, E., Twyman, R., Christou, P. and Fischer, R. (2002) Practical considerations for pharmaceutical antibody production in different crop systems. *Molecular Breeding* **9**, 149-158.

Stoger, E., Vaquero, C., Torres, E., Sack, M., Nicholson, L., Drossard, J., Williams, S., Keen, D., Perrin, Y., Christou, P. and Fischer, R. (2000) Cereal crops as viable production and storage systems for pharmaceutical scFv antibodies. *Plant Molecular Biology* **42**, 583-590.

Stracke, R., Ishihara, H., Barsch, G., Mehrtens, F., Niehaus, K. and Weisshaar, B. (2007) Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the Arabidopsis thaliana seedling. *Plant Journal* **50**, 660-677.

Streatfield, S. (2007) Approaches to achieve high-level heterologous protein production in plants. *Plant Biotechnology Journal* **5**, 2-15.

Streatfield, S. and Howard, J. (2003) Plant-based vaccines. *International Journal For Parasitology* **33**, 479-493.

Streatfield, S., Jilka, J., Hood, E., Turner, D., Bailey, M., Mayor, J., Woodard, S., Beifuss, K., Horn, M., Delaney, D., Tizard, I. and Howard, J. (2001a) Plant-based vaccines: unique advantages. *Vaccine* **19**, 2742-2748.

Streatfield, S.J., Jilka, J.M., Hood, E.E., Turner, D.D., Bailey, M.R., Mayor, J.M., Woodard, S.L., Beifuss, K.K., Horn, M.E. and Delaney, D.E. (2001b) Plant-based vaccines: unique advantages. *Vaccine* **19**, 2742-2748.

Terpe, K. (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology* **60**, 523-533.

Thompson, G. and Larkins, B. (1989) Structural elements regulating zein gene-expression. *Bioessays* **10**, 108-113.

Tremblay, R., Wang, D., Jevnikar, A. and Ma, S. (2010) Tobacco, a highly efficient green bioreactor for production of therapeutic proteins. *Biotechnology Advances* **28**, 214-221.

Turner, R. and Tjian, R. (1989) Leucine repeats and an adjacent dna-binding domain mediate the formation of functional cfos-cjun heterodimers. *Science* **243**, 1689-1694.

Twyman, R., Stoger, E., Schillberg, S., Christou, P. and Fischer, R. (2003) Molecular farming in plants: host systems and expression technology. *Trends in Biotechnology* **21**, 570-578.

Umemura, Y., Ishiduka, T., Yamamoto, R. and Esaka, M. (2004) The Dof domain, a zinc finger DNA-binding domain conserved only in higher plants, truly functions as a Cys2/Cys2 Zn finger domain. *Plant Journal* **37**, 741-749.

Urry, D. (1992) Free-energy transduction in polypeptides and proteins based on inverse temperature transitions. *Progress in Biophysics & Molecular Biology* **57**, 23-57.

Van der Rest, M. and Garrone, R. (1991) Collagen family of proteins. *The FASEB Journal* **5**, 2814-2823.

Van Hamme, J., Singh, A. and Ward, O. (2006) Physiological aspects - Part 1 in a series of papers devoted to surfactants in microbiology and biotechnology. *Biotechnology Advances* **24**, 604-620.

Vaquero, C., Sack, M., Schuster, F., Finnern, R., Drossard, J., Schumann, D., Reimann, A. and Fischer, R. (2002) A carcinoembryonic antigen-specific diabody produced in tobacco. *Faseb Journal* **16**, 408-410.

VicenteCarbajosa, J., Moose, S., Parsons, R. and Schmidt, R. (1997) A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 7685-7690.

Villalobos, A., Ness, J.E., Gustafsson, C., Minshull, J. and Govindarajan, S. (2006) Gene Designer: a synthetic biology tool for constructing artificial DNA segments. *Bmc Bioinformatics* **7**, 285.

Vinson, C., Sigler, P. and Mcknight, S. (1989) Scissors-grip model for dna recognition by a family of leucine zipper proteins. *Science* **246**, 911-916.

Vuorela, A., Myllyharju, J., Nissi, R., Pihlajaniemi, T. and Kivirikko, K. (1997) Assembly of human prolyl 4-hydroxylase and type III collagen in the yeast Pichia pastoris: formation of a stable enzyme tetramer requires coexpression with

collagen and assembly of a stable collagen requires coexpression with prolyl 4-hydroxylase. *Embo Journal* **16**, 6702-6712.

Wilken, L. and Nikolov, Z. (2012) Recovery and purification of plant-made recombinant proteins. *Biotechnology Advances* **30**, 419-433.

Williams, N.A., Hirst, T.R. and Nashar, T.O. (1999) Immune modulation by the cholera-like enterotoxins: From adjuvant to therapeutic. *Immunology Today* **20**, 95-101.

Witte, C.P., Noel, L., Gielbert, J., Parker, J. and Romeis, T. (2004) Rapid one-step protein purification from plant material using the eight-amino acid StrepII epitope. *Plant Molecular Biology* **55**, 135-147.

Wu, J., Yu, L., Li, L., Hu, J., Zhou, J. and Zhou, X. (2007) Oral immunization with transgenic rice seeds expressing VP2 protein of infectious bursal disease virus induces protective immune responses in chickens. *Plant Biotechnology Journal* **5**, 570-578.

Wu, Y., Holding, D. and Messing, J. (2010) gamma-Zeins are essential for endosperm modification in quality protein maize. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 12810-12815.

Wu, Y. and Messing, J. (2012a) Rapid Divergence of Prolamin Gene Promoters of Maize after Gene Amplification and Dispersal. *Genetics* **192,** 507-519.

Wu, Y. and Messing, J. (2012b) RNA Interference Can Rebalance the Nitrogen Sink of Maize Seeds without Losing Hard Endosperm. *Plos One* **7**.

Xie, D., Sharma, S., Wright, E., Wang, Z. and Dixon, R. (2006) Metabolic engineering of proanthocyanidins through co-expression of anthocyanidin reductase and the PAP1 MYB transcription factor. *Plant Journal* **45**, 895-907.

Xu, X., Gan, Q., Clough, R., Pappu, K., Howard, J., Baez, J. and Wang, K. (2011) Hydroxylation of recombinant human collagen type I alpha 1 in transgenic maize co-expressed with a recombinant human prolyl 4-hydroxylase. *Bmc Biotechnology* **11**.

Yanagisawa, S. (2002) The Dof family of plant transcription factors. *Trends in Plant Science* **7**, 555-560.

Yanagisawa, S. (2004) Dof domain proteins: Plant-specific transcription factors associated with diverse phenomena unique to plants. *Plant and Cell Physiology* **45**, 386-391.

Yu, O., Jung, W., Shi, J., Croes, R., Fader, G., McGonigle, B. and Odell, J. (2000) Production of the isoflavones genistein and daidzein in non-legume dicot and monocot tissues. *Plant Physiology* **124**, 781-793.

Yu, O., Shi, J., Hession, A., Maxwell, C., McGonigle, B. and Odell, J. (2003) Metabolic engineering to increase isoflavone biosynthesis in soybean seed. *Phytochemistry* **63**, 753-763.

Zhang, C., Baez, J., Pappu, K. and Glatz, C. (2009a) Purification and Characterization of a Transgenic Corn Grain-Derived Recombinant Collagen Type I Alpha 1. *Biotechnology Progress* **25**, 1660-1668.

Zhang, C., Glatz, C., Fox, S. and Johnson, L. (2009b) Fractionation of Transgenic Corn Seed by Dry and Wet Milling to Recover Recombinant Collagen-Related Proteins. *Biotechnology Progress* **25**, 1396-1401.

Zhu, Z., Hughes, K., Huang, L., Sun, B., Liu, C., Li, Y., Hou, Y. and Li, X. (1994) Expression of human alpha-interferon cdna in transgenic rice plants. *Plant Cell Tissue and Organ Culture* **36**, 197-204.

CHAPTER 3: HYDROXYLATION OF RECOMBINANT HUMAN COLLAGEN TYPE I ALPHA 1 IN TRANSGENIC MAIZE CO-EXPRESSED WITH A RECOMBINANT HUMAN PROLYL 4-HYDROXYLASE

Modified from a paper published in BMC Biotechnology 2011, 11:69

Xing Xu, Qinglei Gan, Richard C. Clough, Kameshwari M. Pappu, John A. Howard, Julio A. Baez, Kan Wang

Abstract

Background Collagens require the hydroxylation of proline (Pro) residues in their triple-helical domain repeating sequence Xaa-Pro-Gly to function properly as a main structural component of the extracellular matrix in animals at physiologically relevant conditions. The regioselective proline hydroxylation is catalyzed by a specific prolyl 4-hydroxylase (P4H) as a posttranslational processing step.

Results A recombinant human collagen type I α-1 (rCIα1) with high percentage of hydroxylated prolines (Hyp) was produced in transgenic maize seeds when co-expressed with both the α- and β- subunits of a recombinant human P4H (rP4H). Germ-specific expression of rCIα1 using maize globulin-1 gene promoter resulted in an average yield of 12 mg/kg seed for the full-length rCIα1 in seeds without co-expression of rP4H and 4 mg/kg seed for the hydroxylated rCIα1 (rCIα1-OH) in seeds with co-expression of rP4H. High-resolution mass spectrometry (HRMS) analysis revealed that nearly half of the collagenous repeating triplets in rCIα1 isolated from rP4H co-expressing maize line had the Pro residues

changed to Hyp residues. The HRMS analysis determined the Hyp content of maize-derived rCIα1-OH as 18.11%, which is comparable to the Hyp level of yeast-derived rCIα1-OH (17.47%) and the native human CIa1 (14.59%), respectively. The increased Hyp percentage was correlated with a markedly enhanced thermal stability of maize-derived rCIα1-OH when compared to the non-hydroxylated rCIα1.

Conclusions This work shows that maize seed has potential to produce adequately modified exogenous proteins with mammalian-like post-translational modifications that may be require for their use as pharmaceutical and industrial products.

Background

Collagen is the most abundant protein found in animals. It has been used widely for industrial and medical applications such as drug delivery and tissue engineering [1, 2]. Human type I collagen is the most abundant collagen type in the human body and is also one of the most studied collagen types. It is a heterotrimer composed of two α1 (CIα1) and one α2 (CIα2) chains with the helical region composed by a repeating composition of Xaa-Yaa-Gly, where X and Y are typically proline (Pro) and hydroxyproline (Hyp) [3]. Collagens used commercially are traditionally extracted from animal tissues. These products contain different types of collagen and may be contaminated with potential immunogenic and infective agents considered hazardous to human health. Thus, recombinant technology has been developed to produce high quality and animal derived contaminant-free collagens. Recombinant collagens have been produced in mammalian cells [4], insect cell cultures [5], yeast [6], and plant cell culture [2, 7].

Transgenic plant systems have advantages over other recombinant production systems in terms of lower cost, higher capacity, lower infective agents/toxins contamination risk, and inexpensive storage capability facilitating processing[8, 9]. The production of plant derived recombinant collagen I α-1 (rCIα1) was reported in 2000 using tobacco [10] and tobacco cell culture [2]. The rCIα1 was also expressed in transgenic maize seed [11, 12] and barley [13].

A challenge for producing rCIα1 in non-mammalian expression systems is the resulting low regioselective hydroxyproline content that makes the product unstable at physiologically relevant temperatures. In humans the 4-hydroxyproline residues synthesized by prolyl 4-hydroxylases (P4Hs) increase the stability of the collagen triple helix structure [14]. The stability of the collagen is increased with the presence of the hydroxyproline primarily through stereoelectronic effects [15]. On the other hand, the hydroxyproline content for the rCIα1 is almost zero in transgenic tobacco[10], or very low in transgenic maize [11] when rCIα1 is not co-expressed with P4H. Since the insect, microbial and plant endogenous P4Hs are not able to achieve the same level of hydroxylation of rCIα1 as present in the human CIα1 chain, the co-expression with collagen of a recombinant animal P4H (rP4H) is necessary to increase the hydroxyproline content of the rCIα1 to deliver a stable product. In tobacco, co-expression of P4H with an α subunit from *C. elegans* and a β subunit from mouse [16] or a recombinant human P4H [17] led to increased hydroxyproline levels of the rCIα1. Similar results were seen in tobacco cell culture [2]. However, the tobacco-derived collagen still had lower Hyp content compared to native human CIα1 making this product unsuitable for use in many applications.

In this study, we generated transgenic maize lines expressing the human rCIα1 gene alone or lines co-expressing both rCIα1 and rP4H genes. Using high-resolution mass spectrometry (HRMS) analysis, we measured the percentages of Hyp and Pro residues in the rCIα1 protein extracted from transgenic maize seeds as well as the actual positions of hydoxylated prolines. We also performed in vitro pepsin treatment at different temperatures to compare the thermal stabilities of maize-derived hydroxylated or non-hydroxylated rCIα1 proteins. Here, we report for the first time that by co-expressing rP4H genes, maize can produce rCIα1 with a hydroxyproline content comparable to native human type I collagen. This achievement provides further confirmation that maize seeds can be used to produce exogenous proteins that require mammalian-like posttranslational modifications for use in specific applications.

Results

Generation of maize lines expressing rCIα1 with and without rP4H co-expression in seeds

The constructs used in this study are shown in Figure 1. The CGB construct carries a gene encoding a recombinant full-length human collagen type I, rCIα1, and the CGD construct carries the rCIα1 gene and both α and β subunits of recombinant human prolyl 4-hydroxylase, rP4Hα and rP4Hβ. The rCIα1 gene was partially maize codon-optimized and its expression was driven by a maize embryo specific globulin-1 promoter (Pglb, [18]). A barley alpha amylase signal sequence (BAASS, [19]) was used as a substitute for the human CIα1 signal peptide (UniProtKB/Swiss-Prot: P02452[1-22],). The combination of embryo specific promoter and the BAASS has demonstrated high expression of foreign proteins in

maize seed [20-22]. The rCIα1 gene lacks the N-propeptide but contains the telopeptide sequences both at the N and C terminal regions. A 29 amino acid bacteriophage T4 fibritin foldon peptide sequence [23] was fused at the C-terminus to the rCIα1 replacing the C-propeptide. The foldon, as the native C-propeptide, facilitates the rCIα1 triple-helical assembly and enhances its stability [23]. To avoid undesired DNA rearrangement caused by using identical sequences (such as using same promoters for multiple gene expression in single construct), we chose to use the maize ubiquitin promoter (Pubi, [24]) to drive the expression of α and β subunits of rP4H. It was shown previously that there is a preferential accumulation of recombinant protein in germ tissue using the ubiquitin promoter[25].

Both constructs were introduced into maize Hi II germplasm using immature embryo via an *Agrobacterium*-based transformation system. Twelve independent transgenic events for CGB and 21 events for CGD, respectively, were recovered and brought to maturity in the greenhouse using pollen donors from an elite inbred. Initial transgene expression screens were conducted on both callus and T1 seeds using an enzyme-linked immunosorbent assay (ELISA) to detect the expression of rCIα1, and α and β subunits of rP4H. T2 seeds from events with the highest transgene expressions were produced by self pollination.

For T1 seed analysis, seeds from multiple plants derived from each event were analysed. In depth molecular and biochemical characterizations of rCIα1 described in this study were performed on T2 seeds from one selected CGB and CGD event, respectively. Individual seeds of the transgenic events were analyzed by polymerase chain reaction (PCR) to separate transgene positive seeds from negative ones. Positive seeds were pooled and analyzed by ELISA for the expression of the rCIα1. Negative null segregant seeds were used as controls.

The average expression level of rCIα1 measured by ELISA in event CGB was 1.86 ± 1.26 mg/g of total soluble protein (TSP) or 12.14 ± 8.06 mg/kg of dry seed weight (DSW). The highest rCIα1 content measured to date from a single CGB seed was 3.54 mg/g TSP or 25.11 mg/kg DSW. The average expression level of the rCIα1 in event CGD was about four times lower than that of in event CGB, which was 0.58 ± 0.26 mg/g TSP or 4.40 ± 2.09 mg/kg DSW. The highest rCIα1 expression in single CGD seed was 0.92 mg/g TSP or 7.54 mg/kg DSW.

Figure 2 shows the detection of rCIα1 in the total protein extracts from CGB and CGD seeds. Because of the low expression level of rCIα1 in CGD seeds, we concentrated the extract using an Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-30 membrane (cat # UFC903008, Millipore) before loading on the gel. Figure 2 shows that rCIα1 could be detected from both CGB and CGD protein samples using anti-foldon antibody. No cross-reacting band at a similar position could be detected in non-transgenic maize seeds (data not shown). It was observed that the CGB rCIα1 migrated faster than its CGD counterpart (Figure 2, open arrow in Lane CGB vs solid arrow in Lane CGD), suggesting different electrophoretic mobility for these two proteins. To exclude that the observed protein migration difference was due to lane shifts during electrophoresis, we mixed the TSPs from both CGB and CGD before loading on the gel. Lane CGB+CGD of Figure 2 shows there are two distinct major bands that cross-reacted with anti-foldon antibody. This result indicates that the rCIα1 proteins derived from maize CGB and CGD events have different electrophoretic mobility, with CGD rCIα1 moves slower than CGB rCIα1. The altered electrophoretic mobility may reflect the increase in molecular weight of rCIα1 that due partially to increased numbers of hydroxylated proline in rCIα1 from CGD event, which is

also co-expressing the rP4H genes. The difference in electrophoretic mobility can also been seen in *Pichia*-derived rCIα1 (Figure 2, FE291) and hydroxylated rCIα1 (Figure 2, FE285).

The expression of the β subunit of rP4H in the CGD seeds was verified by Western Blotting with an anti P4Hβ monoclonal antibody (Figure 3). A main band at ~60 kD (open triangle, Figure 3) was detected in CGD, but not in CGB and non-transgenic wild type maize control, as expected. A weak secondary band detected in CGD is likely due to cross-reactivity of other forms of rP4Hβ in maize. The detection of α subunit of P4H was performed in transgenic callus but not in seeds (data not shown).

To determine whether rCIα1 can also be detected in tissues other than seeds, we performed both protein and transcript analyses of rCIα1 in CGB and CGD plants. Maize leaf samples from 5 different development stages were collected. Total RNA and proteins were prepared from these tissues and subjected to Reverse Transcriptase PCR (RT-PCR) and ELISA, respectively. No detectable rCIα1 transcript and protein could be observed in these samples (data not shown), suggesting that the rCIα1 is not expressed in leaf tissue in both lines as expected.

Co-expression of rP4H increases the hydroxylation of rCIα1

To examine the percentage and positions of the prolines that were hydroxylated by the co-expression of rP4H in the CGD event, we carried out proteomics analysis of gel purified rCIα1 using liquid chromatography tandem mass spectrometry (LC-MS/MS) on the Linear Ion Trap Orbitrap (LTQ Orbitrap) Mass Spectrometer, a high resolution mass spectrometry (HRMS). The HRMS not only can verify the amino acid sequence of the rCIα1, but also can identify the positions of hydroxylated proline residues (Hyp). In addition to maize-derived rCIα1 proteins from CGB and CGD, we also included three control samples:

gel isolated CIα1 fragment from human collagen (cat # 234138, CalBiochem Inc), *Pichia*-derived rCIα1 (isolated from strain FE291 that does not co-express rP4H) and *Pichia*-derived hydroxylated rCIα1 (isolated from strain FE285 that co-express rP4H) (FibroGen).

Results are summarized in Figure 4 and Table 1. The protein sequence coverage by the HRMS (yellow highlighted sequences in Figure 4) on the five samples ranged from 58.66% (human CIα1) to 85.81% (*Pichia* CIα1-OH). To compare the percentages and positions of Hyp in each sample, we chose the peptide regions in all five CIα1 proteins (475 AA) that were covered by the HRMS (red boxes in Figure 4). The common peptide regions represent 44.94% of the full-length CIα1 sequence (1057 AA). A total of 114 Pro and Hyp out of 475 total amino acids (24.00%) were identified by the HRMS in all samples (Figure 4).

For two maize-derived CIα1 samples, a total of 28 and 86 Hyp were identified from CGB and CGD (green highlighted amino acids in Figure 4), respectively, representing a Hyp percentage of 5.89% and 18.11%, respectively, for these two lines (Table 1). This result indicates that the co-expression of rP4H in maize can greatly enhance the hydroxylation of prolines on collagen molecules. The increased number of Hyp in rCIα1 from CGD samples may partially contributed to the increased molecular weight and thereby decreased the migration rate (Figure 2).

Because rP4H catalyzes hydroxylation of Pro residues in the Yaa position of the Xaa-Yaa-Gly triplets within collagen strands [26], we further compared the Pro residues on all Xaa-Pro-Gly triplets in both maize CGB and CGD lines. HRMS analysis identified 752 AA (71.14%) from maize- rCIα1 (CGB) and 818 AA (77.39%) from maize rCIα1-OH (CGD) as shown in Table 1 and Figure 4 (bold and yellow highlighted letters). Among these HRMS identified AA, we chose 652 AA that were shared for both CGB and CGD. We further

identified a total of 90 sets of collagenous triplets within the 652 AA. Among the 90 sets of triplets, 44 sets (48.9%) have the Pro residues changed to Hyp (double underlined triplets in Figure 4) in both CGB and CGD lines; 5 sets (5.6%) have the Pro unchanged (single underlined triplets in Figure 4) in both lines. On the other hand, 41 sets of triplets (45.6%) have the Pro residues changed to Hyp (black boxes in Figure 4) only in rCIα1 isolated from CGD, indicating that nearly half of the collagenous triplets were posttranslationally modified by the co-expression of rP4H genes in CGD maize line.

Seventy-one Hyp residues out of 475 AA by HRMS (14.95%) were identified in human CIα1 control sample (Table 1). For *Pichia* samples, while only two Hyp residues (0.42%) were found in non-hydroxylated CIα1 (FE291), 83 Hyp residues (17.47%) were found in hydroxylated CIα1 (FE285), indicating that the co-expression of P4H in *Pichia* had also dramatically increased proline hydroxylation in collagen (Table 1).

Co-expression of rP4H enhances the thermal stability of rCIα1

To further characterize the maize-derived rCIα1 and rCIα1-OH, we carried out thermal stability analysis using pepsin digestion at 10^{o}C for 15 minutes after heat treatment of protein samples at 4^{o}C or temperatures ranged from 29 to 38.6^{o}C for 6 minutes. The proteolytic resistance of maize-derived collagens were compared with that of the native human collagen and the recombinant collagen from *Pichia pastoris*. Figure 5A is a Western Blot results showing the proteolytic resistance of the collagens after 4°C heat treatment using anti-foldon antibody. Both non-hydroxylated rCIα1 from maize (CGB) and non-hydroxylated rCIα1 from *Pichia* (FE291) were not detected after pepsin treatment. By contrast, the hydroxylated rCIα1 could be detected from both maize (CGD) and

hydroxylated- CIα1 *Pichia* (FE285) samples, suggesting that the pepsin digestion resistance of these collagens was associated with the higher percentage of Hyp residues.

The thermal stability of rCIα1 was further characterized by the determination of melting temperature (Tm) using Western analysis. Two different antibodies, anti-foldon and anti-25kD collagen, were used. In our hands, anti-foldon antibody gave results with less non-specific cross-reactive background bands, while anti-25kD antibody appeared to be more sensitive. Because the native human CIα1 can only be detected with anti-25kD antibody, we used both antibodies in this study. In the experiments shown in Figure 5B, maize seeds TSP from CGB and CGD were extracted and concentrated as described above. The quantity of maize-derived rCIα1 was estimated by ELISA. Approximately 50-100 ng/reaction of rCIα1 from CGB and CGD were used for pepsin treatment. As a control, commercial human collagen (2 μg/reaction) was spiked into TSP extracted from non-transgenic maize seed for pepsin treatment. As can be seen in Figure 5B, both maize-derived rCIα1 (CGB) and rCIα1-OH (CGD) were as stable as the human collagen control at all temperatures tested in the absence of pepsin. When digested with pepsin, the maize-derived non-hydroxylated rCIα1 (CGB) was degraded after the heat treatment at temperatures as low as 4°C. On the other hand, the hydroxylated rCIα1-OH (CGD) could still be detected after temperature treatment as high as 35^{o}C when using anti-foldon antibody, and 38.6°C when using anti-25kD antibody. The difference in Tm results was likely due to the sensitivity and epitope recognition sites of two types of antibodies. Interestingly, the control native human collagen could only withstand the digestion up to temperature treatment around 31°C. This observation is in fact in agreement with the HRMS analysis of the collagens described in Table 1 and Figure 4. Because the maize-derived rCIα1-OH has higher Hyp percentage (18.11%) than that from

human collagen control (14.95%), it is expected that the increased Hyp residues could help to increase the thermal stability of the collagen molecules.

Discussion

The production of plant-derived recombinant collagens have been reported in tobacco leaves, barley cell culture and seeds, as well as maize seeds as summarized in Table 2. Previous tobacco-derived rCIα1 studies showed that different combinations of recombinant human collagens (i.e. rCIα1, rCIα2, and N-propeptide free rCIα1) were used to improve the production of homotrimeric or heterotrimeric recombinant human type I collagen [10, 16, 17, 27]. In a recent paper, Stein et al [17] achieved a high expression level of 200 mg/kg fresh leaves by expressing the collagens under a Chrysanthemum rbcS1 promoter and vacuolar-targeting signal sequence. Early work with tobacco-derived collagens had very low levels of Hyp (0.53%, [10]). With co-expression of C. elegans P4Hα/Mouse P4Hβ [16] or the human rP4Hα/β [17], Hyp levels were increased to 8.41% and 7.55%, respectively. However, this enhanced Hyp level in tobacco is still lower than that of native human collagen CIα1, which is reported as 10.8% by amino acid analysis [28] or around 15% by the HRMS analysis (this work).

Both the full length and a smaller fragment (45 kD) of rCIα1 were produced in barley cell culture [7] and barley seeds [13]. The barley-derived 45 kD collagen has 2.8% of Hyp content when produced in seeds without co-expression of rP4H genes [13].

Previous work on fractionation, purification and characterization of maize-derived full length and a smaller fragment (44 kD) of collagen suggested that an accumulation level of about 3 mg/kg (for the full length) and of 20 mg/kg (for the 44 kD) of DSW, respectively

[11, 29]. A similar maize line accumulating the full length rCIα1 producing maize line (CGB) was used in this study. In our case, the collagen yield of the rCIα1 accumulating line without P4H co-expression averages 12 mg/kg DSW, while the rCIα1 accumulating line with P4H co-expression (CGD) is about 4 mg/kg. The Hyp percentage in rCIα1 protein of CGB was reported as 1.23% using total amino acid composition (AAC) analysis [11]; however, it was measured at 5.89% by using HRMS analysis in our study. Similarly, the Hyp percentage in human CIα1 was reported as 10.8% using AAC analysis [28]. In our HRMS analysis, the Hyp for human CIα1 measured around 15%. It is not clear why Hyp percentages of C1α1 proteins measured uniformly higher in HRMS analysis than that of in AAC analysis. This discrepancy is likely due to the different degrees of resolution of these two very different methodologies. Because the concentrations of rC1α1 and rC1α1-OH obtained from maize seeds were too low to be measured by AAC analysis in our study, we were not able to obtain AAC analysis results for comparison.

P4H is an enzyme that regioselectively modifies the Pro residues in collagenous triplets Xaa-Pro-Gly [30, 31] in the ER as a posttranslational modification. Compared to the *Pichia* recombinant protein production system, maize can produce hydroxylated rCIα1 with a comparable Hyp percentage (Table 1, 18.11% in maize CGD vs 17.47% in *Pichia* FE285). Interestingly, rCIα1 produced in maize seems to have a higher base-level Hyp percentage when compared to rCIα1 isolated from *Pichia* with no rP4H co-expression (Table 1, 5.89% in maize CGB vs 0.42% in *Pichia* FE291). Small numbers of proline at both Xaa and Yaa positions got hydroxylated in CGB maize line without the co-expression of P4H (data not shown). It is likely that the rCIα1 produced in maize is also a substrate for plant endogenous P4Hs with lower efficiency [30].

Conversely, the expression of human rP4H in maize may also catalyze hydroxylation of Pro residues in any plant endogenous proteins with collagenous domains. We checked amino acid sequences of three abundant seed storage proteins (19 kD and 22 kD α-zein, and 27 kD γ-zein) in maize and did not find any collagenous triplets (X. Xu, unpublished). Therefore we do not expect any Pro to Hyp modifications on these seed storage protein in the rP4H expressing CGD line. In fact, the Hyp-only AAC analysis on both CGB and CGD seeds showed no differences in Hyp contents (X. Xu, unpublished). However, because both α and β subunits of rP4H were under the control of the constitutive ubiquitin promoter, it is possible that any of the collagenous triplet domains on proteins in plant cells can be modified by rP4H in such transgenic lines. It may be desirable in the future to restrict rP4H expression to seed tissue only using seed specific promoters.

Using HRMS to analyze posttranslational modification has obvious advantages such as low protein quantity requirement, free of contaminating proteins in samples and reading accuracy. However, it does not give 100% coverage. In this study the peptide coverage ranged from 58.66% to 85.81%.

Because posttranslational modification is a continuous process in the cells, the collagen isolated from the seeds represents a population of protein molecules, i.e., the proline hydroxylation may vary from one collagen molecule to another. In fact, we have performed multiple HRMS measurements on samples extracted from same batch of seeds. We found that while positions of Pro to Hyp modification may vary between measurements, the overall Hyp content remained constant between these samples.

The thermal stability tests in this report showed that maize-derived rCIα1-OH could still be detected after pepsin digestion followed by heat treatment as high as 35^{o}C (using anti-

foldon antibody) and 38.6^{o}C (using anti-25kD antibody). Commercial human collagen control undergoing the same treatment could only withstand up to 31^{o}C temperatures. Stein *et al* [17] reported that the melting temperatures for their tobacco-derived collagen heterotrimer and human skin collagen samples were around 39^{o}C. High melting temperature of plant-derived collagen could potentially be useful for certain industrial application where higher melting temperature is desired, for example, biomaterials for tissue engineering [32, 33].

We recovered the maize-derived rCIα1 from the seed total soluble proteins using a previously described protocol [11]. Because collagens are acid soluble proteins, the extraction buffer used had a pH of 1.7. Unlike Zhang *et al*. [11], we did not perform extensive purification for rCIα1 before gel electrophoresis and Western blot analysis. When treating such acidic rCIα1 solutions under high temperature as we normally do before loading protein gels, we were unable to detect them in Western blot, suggesting that the combination of acidic buffer and high temperature could be detrimental to collagen integrity. Therefore in this study, all maize-derived rCIα1 samples in acidic solutions were not boiled prior to Western blot analysis to avoid collagen degradation.

It is interesting to note that both maize- and Pichia- derived non-hydroxylated rCIα1 were completely digested by pepsin at 10^{o}C after the temperature treatment of samples at 4^{o}C in our study (Figure 5A). This result is different from what reported in barley [7] and maize [11], in which plant- and Pichia-derived rCIα1 were still detectable after the heat treatment around 26-27^{o}C. This could attribute to the different pepsin treatment protocols used in the experiments. For example, the pepsin experiments reported by Zhang et al [11] were conducted under pH 7, 15 minutes heat treatment followed by 150 µg/mL pepsin digestion at

4°C for 16-18 hr. Ritala et al [7] conducted heat treatment for 6 min before subjected the samples to 150 µg/mL pepsin digestion at 10°C for 30 min under an acidic condition. Our conditions were similar to Ritala et al except that we used 200 µg/mL pepsin for 15 min under pH 1.7. Because pepsin functions best in acidic environment, our pepsin digestion under low pH is likely leading to the degradation of non-hydroxylated rCIα1 even at 4°C. Another explanation could be the quantity of the collagen substrate used in different experiments. We estimated that approximately 50-100 ng/reaction of unpurified rCIα1 from CGB seeds were subjected to pepsin digestion in our study. However, Zhang et al used about 600 – 700 ng purified rCIα1 per reaction in their study [11]. The quantity of collagen for pepsin digestion in Ritala et al [7] was not specified.

We have demonstrated for the first time that mammalian-like hydroxylation of human rCIα1 can be achieved in transgenic maize co-expressed with a human rP4H. The Hyp content in maize-derived hydroxylated rCIα1 is comparable to that of the native human version, leading to a similar thermal stability of the product. The current expression levels of collagen reported here are too low for large scale production, as desired accumulation level of recombinant proteins for commercial production is estimated between 250 to 1000 mg/kg grain [34, 35]. Further improvement of recombinant protein production in plants can be achieved by optimization of gene expression including using more effective regulatory elements and protein targeting/retention sequences, as well as using conventional breeding program to select high expression lines over generations [34, 36].

Conclusions

In this study we have shown that properly hydroxylated recombinant human collagen I alpha 1 (rCIα1) can be produced in maize seed. By co-expressing recombinant human prolyl 4-hydroxylases (rP4Hs), we have successfully produced rCIα1 containing Hyp residue levels that are comparable to native human CIα1. The increased Hyp content is associated with increased thermal stability in maize-derived rCIα1. Application of high-resolution mass spectrometry (HRMS) allowed us to measure hydroxylated prolines at specific amino acid positions in different samples. Our findings indicate that maize seed can be used as a system to produce recombinant proteins requiring mammalian-like posttranslational modifications.

Methods

Vector construction

Human collagen type I α 1 (CIα1) coding sequence together with its original N- and C- telopeptides sequences (UniProtKB/Swiss-Prot: P02452) were optimized by Aptagen LLC (Jacobus, PA) for expression in maize. The optimized CIα1 sequence was fused with a 29 amino acids bacteriophage foldon peptide sequence [23] at the C-terminus to produce a protein with 1086 amino acids. Two constructs (Figure 1) were made to produce either recombinant CIα1 (rCIα1) only (CGB), or both rCIα1 and recombinant human prolyl-4-hydroxylase (rP4H, CGD). The rCIα1 gene was regulated by a maize embryo-specific promoter, globulin-1 [18], with a 3'-terminator from potato protease inhibitor II (pin II) gene. Genes encoding two subunits of rP4H (rP4Hα and rP4Hβ) were regulated by a maize constitutive promoter (ubiquitin promoter) and the potato pin II gene terminator. All three gene coding sequences (rCIα1, rP4Hα, and rP4Hβ) in the two constructs were translationally

fused with a barley alpha amylase signal sequence (BAASS, [19]) at the 5' end. The phosphinothricin acetyl transferase (bar) gene driven by the cauliflower mosaic virus (CaMV) 35S promoter was adopted in both constructs to be a marker for the transgenic callus selection. It confers resistance to the herbicide glufosinate ammonium (bialaphos) [37-39].

Production of transgenic plants

Constructs CGB and CGD were introduced into immature embryos of Hi II maize genotype [40] via an *Agrobacterium*-based transformation system [41]. Briefly, maize immature embryos were infected by Agrobacterium strain EHA101 [42] containing the above described vectors and selected on 3 mg/L bialaphos. Regeneration of transgenic plants from the callus was as previously described [20]. Seedlings were transplanted into soil in the greenhouse and allowed to flower and produce seed through hand-pollinations. Seed increases for multiple events from CGB and CGD were conducted in greenhouse and nursery trials. T2 transgenic maize seeds were used for further analysis in this study.

PCR analysis of transgenic plants

Total genomic DNA was isolated by Cetyl Trimethyl Ammonium Bromide (CTAB) method [43] from maize leaf or seed. The presence of transgenes rCIα1, rP4Hα and rP4Hβ were detected by polymerase chain reaction (PCR). A typical PCR reaction consists 100 ng of genomic DNA, 0.8 mM of dNTPs, 2 mM of $MgCl_2$, Taq DNA polymerase buffer and 0.5 U Taq DNA polymerase (Bioline USA Inc, Taunton, MA) in a final volume of 25 µL. PCR was performed at the following condition for 35 cycles: 30 s denaturation at 94°C, 30 s annealing at 60°C, and 45 s extension at 70°C. Primers for amplifying rCIα1 are x7-05 (5'-ACCAGATGGGCCGCTCTCACCTTT-3') and x7-06 (5'-TTCCCTGGTGCCGTTGGAGCTA-3'); for rP4Hα are x7-17 (5'-

ATCTCGGCGTCGCTGATGAT-3') and x7-18 (5'- GTGGTCCGAGCTGGAGAACC -3'); and for rP4Hβ are x7-13 (5'-ATGAAGAACACCTCCTCCCTCTG-3') and x7-14 (5'-TCACAGCTCGTCCTTCACGG -3'). PCR products were analyzed in 1% agarose gel. The expected sizes of PCR products are 1308 bp (rCIα1), 745 bp (rP4Hα) and 1531 bp (rP4Hβ), respectively. Gel was stained by ethidium bromide (0.5 μg/ml) for 20 min. The products size was determined by 1 kb DNA Ladder (cat # N3232S, New England Biolabs).

Protein extraction

Total soluble protein (TSP) from maize seeds was extracted using an acidic buffer described by [11] for collagen preparation. Maize seeds were ground in a coffee grinder (Mr. Coffee) for 1 min. For rCIα1 extraction, extraction buffer (0.1 M phosphoric acid, 0.15 M sodium chloride, pH 1.7) was added in to the seed powder at the ratio of 1:10 (w/v). For rP4H extraction, extraction buffer [25 mM sodium phosphate (pH 6.6), 100 mM sodium chloride, 0.1% Triton X-100 (v/v), 1 mM EDTA, 10 μg/mL leupeptin, and 0.1 mM serine protease inhibitor Perfabloc SC (Fluka)] was added into the seed powder at the ratio of 1:10 (w/v). The mixture was incubated in a shaker incubator (250 rpm, 37°C) for 0.5 hour for rCIα1 and one hour for rP4H. The mixtures were then centrifuged at 13,000 rpm for 10 min at room temperature in a bench top centrifuge. The supernatants were transferred to clean tubes for further analysis. Some protein samples were concentrated by Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-30 membrane (cat # UFC903008, Millipore) followed the product instruction. In short, 15 mL of total seed protein extraction was loaded into the filter device, centrifuged at 3000 x g for approximately 2-3 hours at 4°C. Concentrated samples were recovered by withdrawing with a pipettor. The concentration level was measured by the volume and could be adjusted by the control of the centrifugation time.

ELISA

A competitive ELISA procedure developed by FibroGen and described by Zhang *et al*. [29] was used with minor modifications. Briefly, ELISA plates (cat # 3590, Corning) were coated overnight at 4°C with 5 ng per well of heat-denatured (65± 5°C for 30 minutes) non-hydroxylated rCIα1 from *Pichia pastoris* (FE301,[7]) with phosphate buffer saline (PBS, cat # 21-040-CV, Mediatech). After washing with washing buffer (10 mM PBS, 0.05% Tween 20, pH 7.0), the plates were blocked with 2% dry milk in 100 mM PBS for 1 hour at room temperature. After 3X washings with washing buffer, heat-denatured samples and standard (FE301) in assay buffer (100 mM PBS, 0.05% Tween 20, 1% dry milk, pH 7.0) were added to the plates. The primary antibody, rabbit polyclonal anti-25kDa CIα1 (CA725, FibroGen), was added immediately at a 1:4000 dilution in the assay buffer. After 1 hour incubation at room temperature, plates were washed 3X with washing buffer. The goat-anti-rabbit IgG (H+L) HRP conjugate (cat # 81-6120, Zymed) was added at a 1:5000 dilution in the assay buffer followed by incubation at room temperature for 1 hour. After 3X washings with washing buffer, 100 µL/well of Sure Blue TMB substrate solution (cat # 52-00-01, Kirkegarrd & Perry Laboratories) were added. The plated were then read at 620 nm on a microplate reader (KC4, Biotek) after incubation at room temperature for 30 minutes.

Western blotting

Forty microliters of protein extract from maize seed were mixed with 8 µL of Laemmli sample buffer (cat # 161-0737, Bio-Rad) and then loaded onto a 4-15% polyacrylamide SDS-PAGE gel (cat # 161-1158, Bio-Rad). To avoid protein degradation in the combination of acidic pH and high temperature (X. Xu, unpublished), the step of sample boiling prior to loading was omitted. The proteins separated on the gel were transferred to a

0.45 μm nitrocellulose membrane using Bio-Rad Semidry Transblotting apparatus according to the manufacturer's instructions. Membranes were incubated in blocking buffer (138 mM sodium chloride, 2.7 mM potassium chloride, pH 7.4, 0.1% Tween-20, 5% dry milk powder) for 1 hour at room temperature on a rotary shaker. The membrane was then incubated for 1 hour in blocking buffer with 1:1000 dilution of anti-foldon antibody (rabbit anti-sera with 0.01% sodium azide) for the rCIα1, and with 1:1000 dilution of anti-P4Hβ antibody (cat # 63-164, ICN Biomedicals) for the rP4Hβ. After washing with washing buffer (138 mM sodium chloride, 2.7 mM potassium chloride, pH 7.4, 0.05% Tween-20) 4 times (5 min each wash), the membrane was then incubated for 1 hour in blocking buffer with 1:5000 dilution of HRP-Goat anti-rabbit IgG (H+L) secondary antibody (cat # 62-6120, Zymed) for the rCIα1, and with1:5000 dilution of HRP-Goat anti-mouse IgG (H+L) secondary antibody (cat # 62-6520, Zymed) for the P4Hβ. After washing the membrane with washing buffer 4 x 5 min, the excess buffer was then drained off and the membrane transferred into a clean container. Bands appeared after incubation with horseradish peroxidase substrate, 3,3', 5,5'-tetramethylbenzidine (cat # T0565, Sigma) within 10 minutes.

High-resolution mass spectrometry (HRMS) analysis

To prepare maize-derived rCIα1, 10 μg of total soluble proteins extracted from seeds was separated on the 4-15% polyacrylamide SDS-PAGE gel followed by Bio-Safe Coomassie Stain (cat # 161-0786, Bio-Rad). For purified collagen control samples, three micrograms of each of *Pichia*-derived non-hydroxylated collagen (FE291), *Pichia*-derived hydroxylated collagen (FE285), and human collagen (cat # 234138, CalBiochem Inc.) were loaded on the gel. After electrophoresis, collagen bands were excised from the gels and sent to the Proteomics & Mass Spectrometry Facility at Donald Danforth Plant Science Center, St.

Louis, MO for analysis. The samples were automatically digested with trypsin performed by MultiProbe II protein digester (PerkinElmer) in a temperature-controlled enclosed environment. After digestion, samples were run by LC-MS/MS on the Linear Ion Trap Orbitrap (LTQ-Velos Orbitrap, ThermoFisher Scientific). For post-translational modification analysis, the numbers of Hyp and Pro from each sample were counted and compared.

Thermal stability analysis

The melting temperature (Tm) of C1α1 samples was determined by pepsin digestion after heat treatment [23]. Twenty-five microliters of total soluble protein extracted from CGB and CGD maize seed was subjected to heat treatment in a Thermocycler machine (Biometra GmbH, Germany) at 4^{o}C, or at temperatures ranged from 29°C to 38.6°C for 6 min. For positive controls, 1.4 µg of *Pichia*-derived rCIα1 in hydroxylated (FE285) and non-hydroxylated (FE291) forms, and human collagen were also treated. After heat treatment, all protein samples were then incubated at 10°C with or without pepsin (0.2 mg/mL final concentration, cat # P6887, Sigma) for 15 min. Digestion results were analyzed by western blotting using anti foldon and anti 25 kDa collagen antibodies.

Authors' contributions

XX carried out all the molecular analysis on T_2 transgenic seeds and drafted the manuscript. QG assisted with XX for protein analysis. RC made constructs for maize transformation and conducted the molecular analysis. KP carried out the biochemical analysis in maize plants. JH, JB and KW conceived the study and review the paper. KW designed the experiment and edited the paper. All authors read and approved this final manuscript.

Acknowledgements and Funding

The authors thank Cheng Zhang, Charles E. Glatz, Matt Aspelund and Christopher M Setina for helpful discussions and sharing immune blot materials and methods, Susana Martin Ortigosa for assisting in image preparation, Jessica Zimmer and Angela Nguyen for technical assistances. We also thank Michael Horn, Carol Drees and other former ProdiGene employees for maize transformation, greenhouse care and field operations, and Sheri Almeda and James Polarek for their support to this program at FibroGen. Partial financial support for this project was provided by the United States Department of Agriculture (USDA special grant IOW05082) and the Biopharmaceutical/Bioindustrial Initiative of the Plant Sciences Institute at Iowa State University.

References

1. Lee CH, Singla A, Lee Y: Biomedical applications of collagen. *International Journal of Pharmaceutics* 2001, 221(1-2):1-22.

2. Olsen D, Yang CL, Bodo M, Chang R, Leigh S, Baez J, Carmichael D, Perala M, Hamalainen ER, Jarvinen M *et al*: Recombinant collagen and gelatin for drug delivery. *Advanced Drug Delivery Reviews* 2003, 55(12):1547-1567.

3. Vanderrest M, Garrone R: Collagen Family of Proteins. *FASEB Journal* 1991, 5(13):2814-2823.

4. Geddis AE, Prockop DJ: Expression of human COL1A1 gene in stably transfected HT1080 cells: the production of a thermostable homotrimer of type I collagen in a recombinant system. *Matrix* 1993, 13(5):399-405.

5. Myllyharju J, Lamberg A, Notbohm H, Fietzek PP, Pihlajaniemi T, Kivirikko KI: Expression of wild-type and modified pro alpha chains of human type I procollagen in insect cells leads to the formation of stable [alpha 1(I)](2)alpha 2(I) collagen heterotrimers and [alpha 1(I)](3) homotrimers but not [alpha 2(I)](3) homotrimers. *Journal of Biological Chemistry* 1997, 272(35):21824-21830.

6. Vuorela A, Myllyharju J, Nissi R, Pihlajaniemi T, Kivirikko KI: Assembly of human prolyl 4-hydroxylase and type III collagen in the yeast Pichia pastoris: formation of a stable enzyme tetramer requires coexpression with collagen and assembly of a stable collagen requires coexpression with prolyl 4-hydroxylase. *Embo Journal* 1997, 16(22):6702-6712.

7. Ritala A, Wahlstrom EH, Holkeri H, Hafren A, Makelainen K, Baez J, Makinen K, Nuutila AM: Production of a recombinant industrial protein using barley cell cultures. *Protein Expression and Purification* 2008, 59(2):274-281.

8. Ma JKC, Drake PMW, Christou P: The production of recombinant pharmaceutical proteins in plants. *Nature Reviews Genetics* 2003, 4(10):794-805.

9. Menkhaus TJ, Bai Y, Zhang CM, Nikolov ZL, Glatz CE: Considerations for the recovery of recombinant proteins from plants. *Biotechnology Progress* 2004, 20(4):1001-1014.

10. Ruggiero F, Exposito JY, Bournat P, Gruber V, Perret S, Comte J, Olagnier B, Garrone R, Theisen M: Triple helix assembly and processing of human collagen produced in transgenic tobacco plants. *FEBS Letters* 2000, 469(1):132-136.

11. Zhang C, Baez J, Pappu KM, Glatz CE: Purification and characterization of a transgenic corn grain-derived recombinant collagen type I alpha 1. *Biotechnol Prog* 2009, 25(6):1660-1668.

12. Zhang C, Glatz CE, Fox SR, Johnson LA: Fractionation of Transgenic Corn Seed by Dry and Wet Milling to Recover Recombinant Collagen-Related Proteins. *Biotechnology Progress* 2009, 25(5):1396-1401.

13. Eskelin K, Ritala A, Suntio T, Blumer S, Holkeri H, Wahlstrom EH, Baez J, Makinen K, Maria NA: Production of a recombinant full-length collagen type I alpha-1 and of a 45-kDa collagen type I alpha-1 fragment in barley seeds. *Plant Biotechnol J* 2009, 7(7):657-672.

14. Myllyharju J: Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis. *Matrix Biol* 2003, 22(1):15-24.

15. Kotch FW, Guzei IA, Raines RT: Stabilization of the collagen triple helix by O-methylation of hydroxyproline residues. *J Am Chem Soc* 2008, 130(10):2952-2953.

16. Merle C, Perret S, Lacour T, Jonval V, Hudaverdian S, Garrone R, Ruggiero F, Theisen M: Hydroxylated human homotrimeric collagen I in Agrobacterium tumefaciens-mediated transient expression and in transgenic tobacco plant. *FEBS Lett* 2002, 515(1-3):114-118.

17. Stein H, Wilensky M, Tsafrir Y, Rosenthal M, Amir R, Avraham T, Ofir K, Dgany O, Yayon A, Shoseyov O: Production of bioactive, post-translationally modified, heterotrimeric, human recombinant type-I collagen in transgenic tobacco. *Biomacromolecules* 2009, 10(9):2640-2645.

18. Belanger FC, Kriz AL: Molecular-basis for allelic polymorphism of the maize globulin-1 gene. *Genetics* 1991, 129(3):863-872.

19. Rogers JC: Two barley alpha-amylase gene families are regulated differently in aleurone cells. *The Journal of Biological Chemistry* 1985, 260(6):3731.

20. Hood EE, Bailey MR, Beifuss K, Magallanes-Lundback M, Horn ME, Callaway E, Drees C, Delaney DE, Clough R, Howard JA: Criteria for high-level expression of a fungal laccase gene in transgenic maize. *Plant Biotechnol J* 2003, 1(2):129-140.

21. Hood EE, Love R, Lane J, Bray J, Clough R, Pappu K, Drees C, Hood KR, Yoon S, Ahmad A *et al*: Subcellular targeting is a key condition for high-level accumulation of cellulase protein in transgenic maize seed. *Plant Biotechnology Journal* 2007, 5(6):709-719.

22. Lamphear BJ, Barker DK, Brooks CA, Delaney DE, Lane JR, Beifuss K, Love R, Thompson K, Mayor J, Clough R *et al*: Expression of the sweet protein brazzein in maize for production of a new commercial sweetener. *Plant Biotechnol J* 2005, 3(1):103-114.

23. Pakkanen O, Hamalainen ER, Kivirikko KI, Myllyharju J: Assembly of stable human type I and III collagen molecules from hydroxylated recombinant chains in the yeast Pichia pastoris - Effect of an engineered C-terminal oligomerization domain foldon. *Journal of Biological Chemistry* 2003, 278(34):32478-32483.

24. Christensen AH, Sharrock RA, Quail PH: Maize polyubiquitin genes - structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Molecular Biology* 1992, 18(4):675-689.

25. Witcher DR, Hood EE, Peterson D, Bailey M, Bond D, Kusnadi A, Evangelista R, Nikolov Z, Wooge C, Mehigh R *et al*: Commercial production of beta-glucuronidase (GUS): a model system for the production of proteins in plants. *Molecular Breeding* 1998, 4(4):301-312.

26. Hutton JJ, Jr., Kaplan A, Udenfriend S: Conversion of the amino acid sequence gly-pro-pro in protein to gly-pro-hyp by collagen proline hydroxylase. *Arch Biochem Biophys* 1967, 121(2):384-391.

27. Perret S, Merle C, Bernocco S, Berland P, Garrone R, Hulmes DJ, Theisen M, Ruggiero F: Unhydroxylated triple helical collagen I produced in transgenic plants provides new clues on the role of hydroxyproline in collagen folding and fibril formation. *J Biol Chem* 2001, 276(47):43693-43698.

28. Miller EJ, Gay S: Collagen: An overview. *Methods Enzymol* 1982, 82 Pt A:3-32.

29. Zhang C, Baez J, Glatz CE: Purification and characterization of a 44-kDa recombinant collagen I alpha 1 fragment from corn grain. *J Agric Food Chem* 2009, 57(3):880-887.

30. Gorres KL, Raines RT: Prolyl 4-hydroxylase. *Critical Reviews in Biochemistry and Molecular Biology* 2010, 45(2):106-124.

31. Kivirikko KI, Myllylä, R. and Pihlajaniemi, T: Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins. In: *Post-translational modifications of proteins.* Edited by Harding JJC, M. James C. Boca Raton: CRC Press; 1992: 1-52.

32. Shoulders MD, Raines RT: Collagen Structure and Stability. *Annual Review of Biochemistry* 2009, 78:929-958.

33. Yang C, Hillas PJ, Baez JA, Nokelainen M, Balan J, Tang J, Spiro R, Polarek JW: The application of recombinant human collagen in tissue engineering. *BioDrugs* 2004, 18(2):103-119.

34. Hood EE, Witcher DR, Maddock S, Meyer T, Baszczynski C, Bailey M, Flynn P, Register J, Marshall L, Bond D *et al*: Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. *Molecular Breeding* 1997, 3(4):291-306.

35. Nandi S, Yalda D, Lu S, Nikolov Z, Misaki R, Fujiyama K, Huang N: Process development and economic evaluation of recombinant human lactoferrin expressed in rice grain. *Transgenic Research* 2005, 14(3):237-249.

36. Chikwamba R, McMurray J, Shou HX, Frame B, Pegg SE, Scott P, Mason H, Wang K: Expression of a synthetic E-coli heat-labile enterotoxin B sub-unit (LT-B) in maize. *Molecular Breeding* 2002, 10(4):253-265.

37. Anzai H: Transgenic tobacco resistant to a bacterial disease by the detoxification of a pathogenic toxin. *Molecular & General Genetics* 1989, 219(3):492.

38. Gordon-Kamm WJ: Transformation of maize cells and regeneration of fertile transgenic plants. *Plant Cell* 1990, 2(7):603.

39. Uchimiya H: Bialaphos treatment of transgenic rice plants expressing a bar gene prevents infection by the sheath blight pathogen (Rhizoctonia solani). *Nature Biotechnology* 1993, 11(7):835.

40. Armstrong, C.L., Green, C.E. and Phillips, R.L: Development and availability of germplasm with high Type II culture formation response. *Maize Genet Coop Newsl*. 1991, 65, 92-93.

41. Ishida Y, Saito H, Ohta S, Hiei Y, Komari T, Kumashiro T: High efficiency transformation of maize (Zea mays L) mediated by Agrobacterium tumefaciens. *Nature Biotechnology* 1996, 14(6):745-750.

42. Hood EE, Helmer GL, Fraley RT, Chilton MD: The Hypervirulence of Agrobacterium-Tumefaciens A281 Is Encoded in a Region of Ptibo542 Outside of Transfer DNA. *Journal of Bacteriology* 1986, 168(3):1291-1301.

43. Murray MG, Thompson WF: Rapid Isolation of High Molecular-Weight Plant DNA. *Nucleic Acids Research* 1980, 8(19):4321-4325.

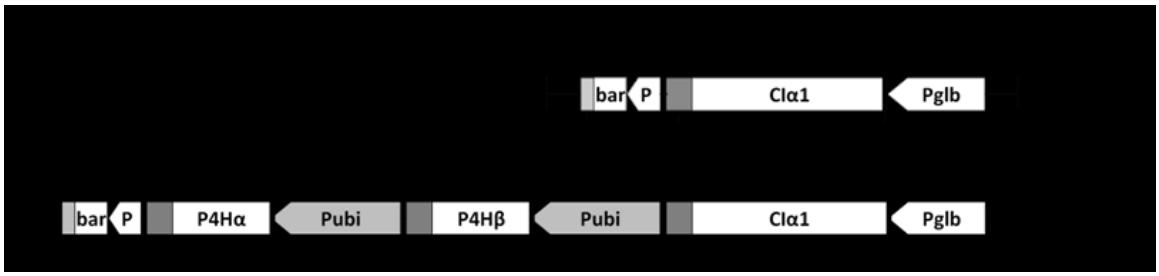

Figure 1. **A schematic representation of the two constructs used in this study**. LB, left border of *Agrobacterium* T-DNA; T35S, CaMV 35S terminator; bar, bialaphos resistant coding sequence; P, CaMV 35S promoter; PIN II, potato protease inhibitor II gene terminator; CIα1, human collagen I α1 chain coding sequence; BAASS, barley alpha amylase signal sequence; P4Hα, prolyl 4-hydroxylase α subunit; P4Hβ, prolyl 4-hydroxylase β subunit; Pglb, maize globulin-1 promoter; Pubi, maize ubiquitin promoter; RB, right border of *Agrobacterium* T-DNA.

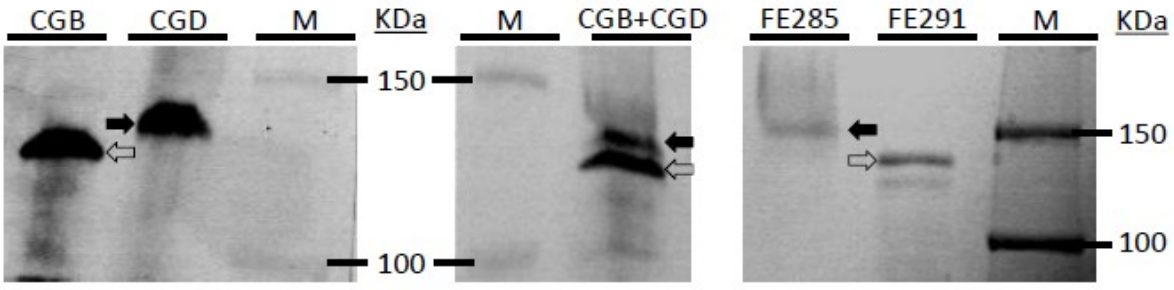

Figure 2. Electrophoretic mobility difference analysis of rCIα1 in the CGB and CGD line. Equal volumes of total protein extracts from seeds of CGB, CGD (10X concentrated by volume) and mixture of CGB+CGD extracts were loaded on the gel. The rCIα1 from CGB and CGD lines were detected by western blot using anti-foldon antibody. *Pichia*-derived rCIα1 (FE291) and rCIα1-OH (FE285) were included as controls and detected by Coomassie Brilliant Blue staining. Open arrows, rCIα1 from CGB or FE291; solid arrows, OH-rCIα1 from CGD or FE285. M, molecular weight marker.

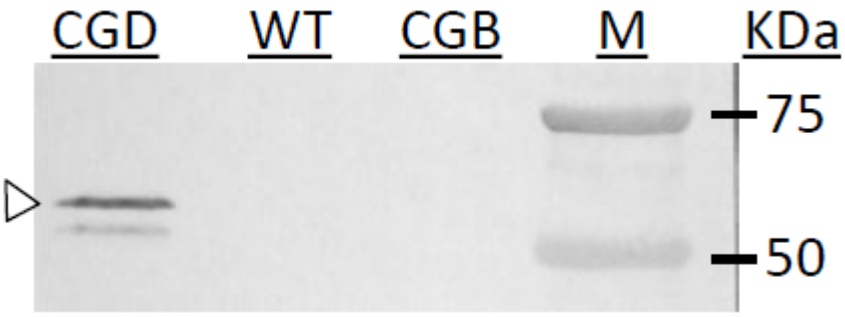

Figure 3. Western blot of the rP4Hβ using anti-P4Hβ antibody. Equal volume of total protein extraction from seeds of CGB, CGD and non-transgenic control maize (WT) extracts was loaded on the gel. Open triangle, rP4Hβ. M: molecular weight marker.

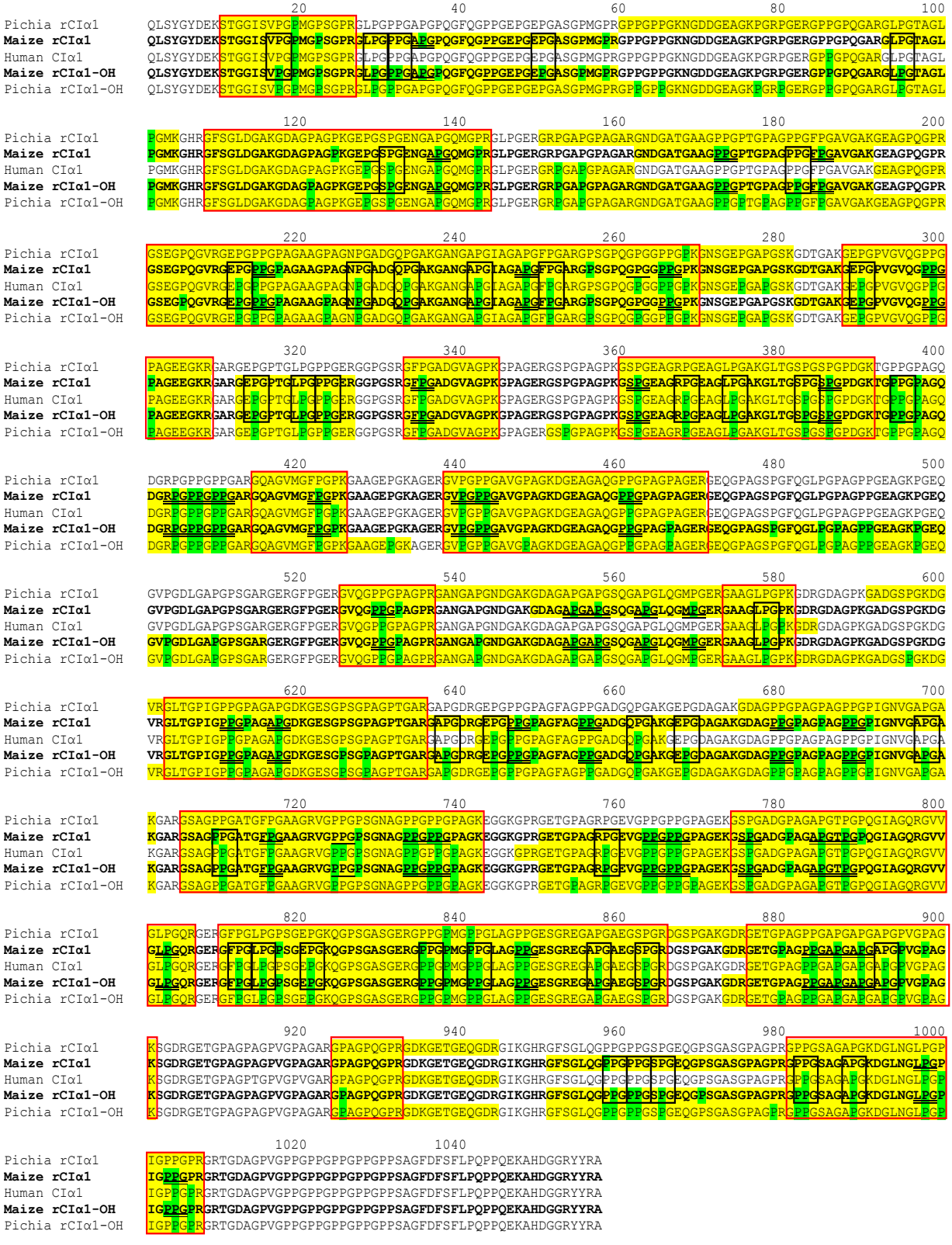
Pichia rCIα1
Maize rCIα1
Human CIα1
Maize rCIα1-OH
Pichia rCIα1-OH

Figure 4. LC-MS/MS analysis of the rCIα1. Full length peptide sequences of 1057 amino acid are listed. Pichia rCIα1, *Pichia*-derived rCIα1 from strain FE291; Maize rCIα1, maize-derived rCIα1 from line CGB; Human CIα1, gel-isolated CIα1 fragment from commercial collagen (CalBiochem Inc); Maize rCIα1-OH, maize-derived rCIα1 from line CGD; Pichia rCIα1-OH, *Pichia*-derived rCIα1 from strain FE285. Yellow-highlighted letters: amino acid sequences identified by the Orbitrap; green-highlighted letters: Hyp residues identified by the Orbitrap; red boxes: peptide regions identified in all five samples by the Orbitrap. Black boxes: collagenous triplets Xaa-Pro-Gly with Pro changed to Hyp in Maize rCIα1-OH but not in Maize rCIα1; single underlines: triplets with Pro unchanged in both maize lines; double underlines: triplets with Pro changed to Hyp in both maize lines.

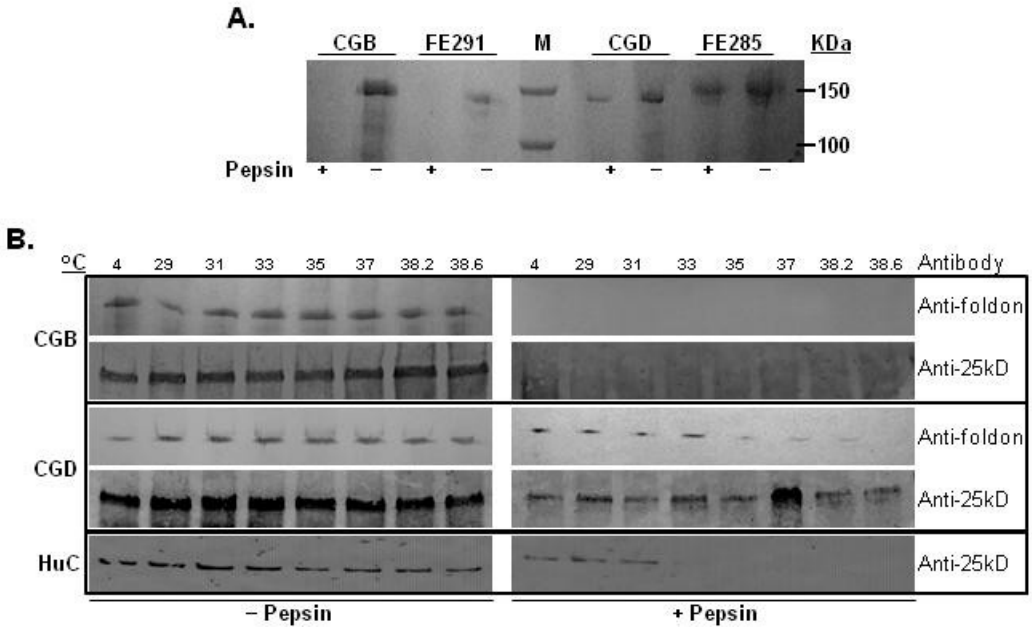

Figure 5. Thermal stability analysis of the rCIα1 from maize, Pichia and human. (A) Western blot results of the maize-derived rCIα1 (CGB), rCIα1-OH (CGD), *Pichia*-derived rCIα1 (FE291), and rCIα1-OH (FE285) after 4°C incubation and pepsin treatment. (B) Western blot results of the maize-derived rCIα1 (CGB), rCIα1-OH (CGD), and human CIα1 (HuC) after heat treatments under various temperatures and pepsin treatments as indicated. M: molecular weight marker.

Table 1. Summary of HRMS analysis on five C1α1 (1057 AA[1]) samples from Maize, *Pichia* and Human

	Pichia Clα1 (FE291)	Maize Clα1 (CGB)	Human Clα1	Maize Clα1-OH (CGD)	Pichia Clα1-OH (FE285)
Peptides identified by HRMS (highlighted in yellow, Figure 4)					
Total # AA	680	752	620	818	907
Percent HRMS coverage	64.33%	71.14%	58.66%	77.39%	85.81%
Peptide regions identified in all five C1α1 (475 AA) by HRMS (red boxes, Figure 4)					
Total # HYP identified by HRMS, in green	2	28	71	86	83
Percent HYP identified by HRMS	0.42%	5.89%	14.95%	18.11%	17.47%
% HYP (by AA analysis)	N/A	1.23%[2]	10.8%[3]	N/A	11.54%[2]

[1] sequences presents in all five samples

[2] from refference 11

[3] from refference 28

Table 2 Summary of plant derived recombinant human collagen I α 1

Expression system	Collagen			rP4H	Hydroxylation content (%)	Reference
	Regulatory sequences	Gene	Yield			
Tobacco Leaf	P35S (constitutive) + PR-protein SP	proα1(I) ΔNproα1(I)	30 mg/kg powdered plants	N/A	0.53	10
Tobacco Leaf	P35S (constitutive) + PR-protein SP	ΔNproα1(I)	N/A	N/A	N/A	27
Tobacco Leaf	L3 + PR-protein SP	ΔNproα1(I)	0.5-1 mg/kg leaf material	P1287 + Native SP + C. elegans P4Hα/Mouse P4Hβ	8.41	16
Tobacco	rbcS1 (constitutive) + vacuole or apoplast targeting SPs	proα1(I)/ proα2(I)	200 mg/kg fresh leaves	P35S (constitutive) + vacuole or apoplast targeting SPs + Human P4Hα/β	7.55	17
Barley P1 cell	Ubi (constitutive) + At chitinase SP + HDEL (ER retention)	proα1(I)	2-9 μg/L cell culture	N/A	N/A	7
Barley Seed	GluB1 (endosperm specific) + At chitinase SP + HDEL (ER retention)	Clα1 45 kD	Below detectable level (Clα1) 45 mg/kg seed (45 kD)	N/A	N/A (Clα1) 2.8 (45 kD)	13
Maize Seed	globulin-1 (embryo specific) + barley α-amylase SP	44 kD	20 mg/kg seed	N/A	2.01	29
Maize Seed	globulin-1 (embryo specific) + barley α-amylase SP	Clα1	3 mg/kg seed	N/A	1.23	11
Maize Seed	globulin-1 (embryo specific) + barley α-amylase SP	Clα1 44 kD	15.9 mg/kg germ (Clα1) 49.6 mg/kg germ (44 kD)	N/A	N/A	12
Maize Seed	globulin-1 (embryo specific) + barley α-amylase SP	Clα1	12 mg/kg seed (Clα1) 4 mg/kg seed (Clα1-OH)	Pubi (constitutive) + Barley α-amylase SP + Human P4Hα/β	18.11	This study

proα1(I): human type I procollagen α 1 chain

ΔproαΔ1(I): human type I procollagen α 1 chain lacking N-propeptide

ΔNCproα1(I): human type I procollagen α 1 chain lacking N-propeptide and C-propeptide

Clα1: sequence information is not clear

44 kD: 44 kD fragment of Clα1

45 kD: 45 kD fragment of Clα1

CHAPTER 4: STRATEGIES AND METHODS OF PRODUCING A DESIGNED PEPTIDE SURFACTANT IN TRANSGENIC MAIZE

Xing Xu, Yan Jin, and Kan Wang

Abstract

Designed peptide surfactants are sustainable, environmentally-friendly alternatives to petroleum-based products within a wide range of applications, including biomedical, pharmaceutical, environmental, and food industries. Currently, these peptide surfactants have been produced mainly in the microbial systems. Because maize seed has the potential to produce low-cost, high-capacity, and low-contamination-risk peptide biosurfactants, we designed strategies to express a designed 22-amino-acid peptide biosurfactant GAM1 in transgenic maize. Our strategy is to first recover the recombinant peptide GAM1 from the endosperm of seed, then use the seed remnant for production of fuel ethanol. GAM1 was translationally fused to the C-terminus of a red fluorescent protein (RFP) or 22-kDa maize α-zein protein. Transgenic maize lines were successfully produced. DNA sequence analysis on transgenic plants showed both translationally-fused genes were correctly assembled. Red fluorescence was detected from seeds expressing RFP::GAM1, suggesting RFP was folded properly in the tissue. However, the presence of GAM1 in RFP::GAM1 fusion could not be detected by using Western analysis and in-gel digestion/peptide mass fingerprinting. Similarly, high-performance liquid chromatography and mass spectrometry failed to detect the presence of GAM1 in Zein::GAM1 fusion. Further experiments are needed to determine whether the failure of

GAM1 detection was due to the low expression of the fusion proteins, degradation of fusion proteins in the process, or other unexpected causes.

Introduction

With the increased demand for fuels and environmental concerns regarding fossil fuels, bio-renewable resources, especially plant-based bio-refineries, are becoming an increasingly important research area. Bio-renewable chemicals target a slightly smaller market, but of higher value with less sensitivity to oil prices. Biosurfactants particularly have been utilized and explored in recent years as alternatives to petroleum-based surfactants (Banat et al., 2000; Clapes and Infante, 2002; Kaar et al., 2009; Sekhon, 2006; Van Hamme et al., 2006). Biosurfactants are bio-derived, surface-active substances with abilities to reduce surfaces and tension at interfaces between two liquids, or between a liquid and a solid (Banat et al., 2000). Biosurfactants have been applied widely on biomedical, pharmaceutical, environmental, and food industries (Banat et al., 2010; Nitschke and Costa, 2007; Singh et al., 2007).

Interfacially-active designed peptide biosurfactants have been recently reported to have advantages of lowering interfacial tension and controling interfacial rheology (Dexter et al., 2006; Dexter and Middelberg, 2007; Malcolm et al., 2006). Bacterial derived lipopeptide biosurfactants (e.g. surfactin) have received much attention due to their high surface activities (Deleu et al., 1999; Grangemard et al., 1997). However, recombinant production of such lipopeptides is inefficient because multiple enzymes are needed in the process (Peypoux et al., 1999). Designed peptide biosurfactants are linear amino acid sequences without conjugated lipids and their production is simpler than for

bacterial lipopeptides (Kaar et al., 2009). The peptide sequences can be designed to provide reversible control of emulsions and foams (Dexter et al., 2006; Malcolm et al., 2006). Peptide biosurfactants are not only designed for material properties, but also can be adopted as antimicrobials (Vogt and Bechinger, 1999), vaccines (Kanodia and Kast, 2008), gene delivery vehicles (Kichler et al., 2003), and nutraceuticals (Mine and Shahidi, 2006).

In the present study, a "switchable" biosurfactant GAM1, slightly modified from a previously reported stimuli-responsive biosurfactant AM1, was chosen as a peptide model (Dexter et al., 2006; Dexter and Middelberg, 2008; Kaar et al., 2009). The AM1 is a 21- residue amphiphilic peptide, which switches the mechanical strength of an oil-water interface according to the aqueous solution composition (Dexter et al., 2006). The favored condition for the helical conformation of AM1 is neutral pH (Dexter et al., 2006). In the presence of zinc ions at neutral pH, the AM1 forms a strong interfacial film caused by the ion-mediated cross-linking between peptides, which stabilizes foams and emulsions (Dexter et al., 2006). At low pH or without metal ions, the surface activity of AM1 is lost; therefore, it is called "switchable" (Dexter et al., 2006). The AM1 is highly soluble in aqueous solution and has no critical micellar behavior (Dexter et al., 2006). The GAM1 has the same properties as the AM1, but with an extra amino acid, glycine, at the N-terminus (Kaar et al., 2009)—a 2489 Dalton, 22-residue amphiphilic peptide (sequence GMKQLADSLHQLARQVSRLEHA).

The biological production of designed peptides are mainly reported for *E. coli* (Achmuller et al., 2007; Kaar et al., 2009; Lee et al., 2005; Morreale et al., 2004; Xu et al., 2006; Zhou et al., 2007), as fusion proteins because of the low expression,

degradation, (Itakura et al., 1977) or toxicity to the host (Ingham and Moore, 2007). Transgenic plants have advantages over other recombinant protein/peptide expression systems, including cost efficiency, higher capacity, lower toxic contamination risk, and inexpensive storage capability (Ma et al., 2003; Menkhaus et al., 2004). Maize seeds have been considered to be a suitable recombinant protein expression system (Azzoni et al., 2002; Bailey et al., 2004; Ramessar et al., 2008; Woodard et al., 2003; Xu et al., 2011). After recovering recombinant proteins from seeds, typically 99% of the seed mass remains (Paraman et al., 2010). The seed remnant can be used for other purposes, such as producing ethanol (Johnson et al., 2003). Thus, the co-production strategy of expressing seed specific recombinant proteins and using the seed remnant after protein extraction for producing fuel ethanol would efficiently utilize the biomass of maize grains and reduce total costs. Additional benefits of using this strategy include improving the environment by reducing waste, lowering green house gas emissions and reducing the U.S. reliance on foreign fossil fuels.

In the present study, we generated transgenic maize lines producing GAM1 by using a protein fusion strategy. GAM1 is a 22-amino acid long peptide. It is difficult to produce the small peptide in *E. coli* (Piers et al., 1993) because small peptides are susceptible to degradation (Li, 2009). The fusion protein strategy has been successfully developed to overcome this challenge. In *E. coli*, GAM1 was fused to the N-terminus of maltose binding protein (MBP), which enhanced its solubility (Kapust and Waugh, 1999) and expression level (Kaar et al., 2009). The high expression of small peptides in transgenic plants has been achieved by fusing to β-glucuronidase (GUS) or fusing two small peptides via a linker sequence (Francois et al., 2002; Okamoto et al., 1998).

The first strategy was to fuse the GAM1 at the C-terminus of the RFP, a commonly used fluorescent tag protein with various commercial purification kits. The second strategy was to fuse GAM1 at the C-terminus of a 22-kDa α-zein, one of the major maize seed storage proteins (Soave et al., 1981). We expected fusing to the 22-kDa α-zein and using the native 22-kDa α-zein promoter and terminator, the GAM1 would achieve high expression as it became a part of the major maize endogenous protein. This strategy has been demonstrated by Dr. David Jackson, Cold Spring Harbor Laboratory (D. Jackson, personal communication). The tobacco etch virus (TEV) protease recognition site was designed between GAM1 and the fusion protein partners, which facilitates the release of GAM1 by enzymatic digestion (Dougherty et al., 1988). Our hypothesis was, by fusing to a tag protein or a major seed protein, we will be able to produce the fusion proteins in seeds and extract GAM1 by using an affinity column (RFP::GAM1) or by size differentiation (zein::GAM1).

Materials and Methods

Vector construction

Two constructs were designed to deliver GAM1 into maize—pYJ16 and pYJ15 (Figure 1). The binary vector pTF101.1 (Paz et al., 2004) was used as a backbone for both constructs provided by the Plant Transformation Facility at Iowa State University. In pYJ16, the GAM1 was fused to the N-terminus of the RFP (RFP::GAM1) with an AG linker sequence (Ala-Gly-Ala-Gly-Ala-Gly) and TEV protease recognition site (Glu-Asn-Leu-Tyr-Phe-Gln-Gly) in between. The AG linker connects two proteins with minimal effects on stability and folding rates of proteins (Ladurner and Fersht, 1997). The TEV

protease recognition site was cut by the TEV protease and the GMA1 was released from the fusion proteins (Dougherty et al., 1988). The signal peptide of a maize seed storage protein—27-kDa γ-zein was added at the N-terminus of the RFP to lead the fusion protein to the reticulum endoplasmic (ER). An ER retention sequence SEKDEL was added after the GAM1 to retain the protein in ER (Gomord et al., 1997). The RFP::GAM1 was driven by the promoter of 27-kDa γ-zein. The terminator in pYJ16 was a soybean vegetative storage protein gene terminator (Tvsp).

In pYJ15, the GAM1 was fused to the C-terminus of another maize seed storage protein, 22-kDa α-zein (22-kDa α-zein::GAM1), also with the AG linker sequence and TEV protease recognition site in between. The promoter of 22-kDa α-zein was used to drive the α-zein::GAM1 in pYJ15. The native terminator of 22-kDa α-zein was also used for pYJ15. The phosphinothricin acetyl transferase (bar) gene driven by the cauliflower mosaic virus (CaMV) 35S promoter was used in both constructs as the marker for the transgenic callus selection (Anzai et al., 1989; Gordonkamm et al., 1990; Uchimiya et al., 1993). *Agrobacterium tumefaciens* strain EHA101 was used to transform both constructs into maize (Hood et al., 1986),

Production of transgenic plants

The maize transformation was performed by the Plant Transformation Facility at Iowa State University as previously described (Frame et al., 2002). Briefly, constructs pYJ15 and pYJ16 were introduced into immature embryos of Hi II maize genotype via an *Agrobacterium*-based transformation system. Herbicide-resistant calli were analyzed by the polymerase chain reaction (PCR) for presence of intact transgenes. Selected calli

were regenerated. The plants were brought to maturity in the Plant Transformation Facility's greenhouse. The transgenic lines were named A273 and A274 for constructs pYJ15 and pYJ16, respectively.

PCR analysis of transgenic plants

Total genomic DNA was isolated using the cetyl trimethyl ammonium bromide (CTAB) method (Murray and Thompson, 1980) from maize calli, leaves, or seed. The presence of the T-DNA was detected by PCR. A PCR reaction contains 100 ng of genomic DNA, 0.8 mM of dNTPs, 2 mM of $MgCl_2$, 1 X Taq DNA polymerase buffer and 0.5 U Taq DNA polymerase (Bioline USA Inc, Taunton, MA) in a final volume of 25 µL. PCR was performed utilizing the following conditions for analysis. For A274 (transformed with pYJ16): 35 cycles by primers x10-09 (5'- CCTACAAGACCGACATC AAG-3') and x10-10 (5'- GTGCATATCAGCATACCTTA-3') (Figure 1): 30 s denaturation at 95°C, 30 s annealing at 56°C, and 20 s extension at 70°C. For A274 (transformed with pYJ15), PCR was performed by utilizing the following conditions by primers x11-01 (5'-TTACAACAGTTGCTTCCATT-3') and x11-02 (5'-TCATTAGGCACCCCA-3') (Figure 1) for 35 cycles—30 s denaturation at 95°C, 30 s annealing at 58°C, and 40 s extension at 70°C. PCR products were analyzed in 1% agarose gel. The expected sizes of PCR products were 355 bp (pYJ16) and 1566 bp (pYJ15). The gel was stained with ethidium bromide (0.5 µg/ml) for 20 min. The product's size was determined by 1 kb DNA Ladder (cat # N3232S, New England Biolabs).

Protein extraction

Total zein proteins (water-insoluble) were extracted from maize seed by using an ethanol-based extraction buffer containing 70% ethanol, 61 mM sodium acetate, and 100 nm DL-Dithiothreitol (DTT) (Flint-Garcia et al., 2009). Seeds were ground and 100 mg the powder was extracted with 1 mL of extraction buffer in a 1.5 mL Eppendorf tube. The tube was laid horizontally on a shaker at 225 rpm at 37^{o}C for 1 hour. Then, the tube was spun 10 minutes at room temperature in a micro-centrifuge at 15,000 rpm. The supernatant was filtered for high-performance liquid chromatography (HPLC), by using Ultrafree centrifugal filter units (cat # UFC30HV00, Millipore) by following the manufacturer's instructions.

Total water soluble protein (TSP) from ground maize seeds was extracted, utilizing the following protein extraction buffer at a ratio of 10 µL buffer per milligram of seed powder for 2 hours at 37^{o}C: 25 mM sodium phosphate (pH 6.6), 100 mM NaCl, 0.1% Triton X-100 (v/v), 1 mM EDTA, 10 µg/ml of leupeptin, and 0.1 mM serine protease inhibitor Perfabloc SC (Fluka).

Fluorescence imaging

The mature seeds of transgenic line A274 were visualized and imaged by utilizing an Olympus SHZ10 stereoscope (Leeds Precision Instruments, Inc., Minneapolis, MN, USA) coupled to a SPOT RT color CCD camera (Diagnostic Instrument Inc., Sterling Heights, MI). Images were taken under a bright or fluorescent field. A RFP filter (Cat # 49008, ET - mCherry, Texas Red, Chroma Technology Corp) was utilized and λ_{ex}=540-580 nm and λ_{em}=600-660 nm were acquired by employing SPOT Advanced software.

Western Blot and Coomassie Brilliant Blue staining

The Red Fluorescent Protein (RFP) from the total soluble protein extracted from A274 seed powder was purified by Anti-RFP magnetic beads (cat # M165-9, MBL International), following the manufacturer's protocol. After affinity column purification, 30 µl of elutes in Laemmli sample buffer were obtained from the 500 µl of A274 raw protein extract. Fifteen micro liters of the A274 samples and 1 µg of the commercial RFP control (cat # 4997-100 BioVision) were loaded onto a 12% polyacrylamide SDS-PAGE gel (cat # 465-1043, Bio-Rad). Electrophoresis was performed at 200 V for 35 min. The gel was stained by Bio-Safe Coomassie Stain (cat # 161-0786, Bio-Rad) or transferred to a 0.45 µm nitrocellulose membrane for the Western Blot using a Bio-Rad Semidry Trans-blotting apparatus, according to the manufacturer's instructions. The membrane was incubated in the blocking buffer (138 mM sodium chloride, 2.7 mM potassium chloride, pH 7.4, 0.1% Tween-20, 5% dry milk powder) for 1 hour at room temperature on a rotary shaker. The membrane was then incubated for 1 hour in the blocking buffer with 1:1000 dilution of rabbit anti-RFP antibody (cat # A00682 Genscript). After washing with the washing buffer (138 mM sodium chloride, 2.7 mM potassium chloride, pH 7.4, 0.05% Tween-20) 4 times (5 min each wash), the membrane was then incubated for 1 hour in the blocking buffer with 1:2000 dilution of HRPGoat anti-rabbit IgG (H+L) secondary antibody (cat # 62-6120, Zymed). The membrane was then washed with the washing buffer 4 × 5 min. The excess buffer was drained and the membrane transferred into a clean container. Bands appeared after incubation with horseradish peroxidase substrate, 3,3', 5,5'-tetramethylbenzidine (cat # T0565, Sigma) within 10 min.

The zein Western Blot procedure is similar to the RFP, except the first antibody was an anti-zein rabbit antibody (kindly offered by Dr. Paul Scott) with the dilution of 1:50,000 instead.

TEV treatment

The TEV (tobacco etch virus) protease treatment is performed under the manufacturer's instruction (cat # 12575-015, Invitrogen). In short, 50 µL of the RFP affinity column purified protein sample from seed for event A274-9-2 was digested in a solution containing 7.5 µL 20X TEV buffer (1 mM Tris-HCl pH 8.0, 10 mM EDTA), 0.1 M DTT (dithiothreitol), 10 units (1 µL) AcTEV^{TM} Protease (cat # 12575-015, Invitrogen), and 90 µL water. The reaction was performed at 30 °C for 2 hours. The reaction can be terminated by adding Laemmli sample buffer and stored at -20 °C.

High-performance liquid chromatography (HPLC) and mass spectrometry

The HPLC method for α-zein::GAM1 was slightly modified from Flint-Garcia et al. (2009). In brief, the alcohol-soluble zein proteins were extracted as described in protein extraction. An aliquot (20 µL) of the extract was injected into a C18 column (5 microns, 2.1 mm X 250 mm, Vydac Cat # 218TP52), and absorbance at 214 nm was measured. Separation of zein proteins was achieved by hydrophobicity with a gradient of water and acetonitrile. The water contained 0.1 % trifluoroacetic acid and the acetonitrile contained 0.08 % trifluoroacetic acid. The gradient of acetonitrile ranged as follows—0-30 % from time 0 to 20 min, 30-60 % from time 20 to 28 min, 60-63 % from time 28 to 40 min, and 63-100 % from time 40 to 45 min. The flow rate was 0.2 mL/min. All peaks

from the HPLC were collected individually and analyzed further by Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). The Protein Facility at Iowa State University completed the HPLC and MS work.

The HPLC method for RFP::GAM1 was provided by Dr. Charles E. Glatz (Department of Chemical and Biological Engineering, Iowa State University). An aliquot of 20 μL of the affinity-column-purified RFP was injected into a Phenomenex Jupiter C-5 column (5 u particle size, 300 Angstrom Unit pore size, 4.6 mm x 150 mm). Mobile phases are: A) 0.1% trifluoroacetic acid and B) 90% acetonitrile. The flow rate was 1 mL/min. and the gradient was 25 to 55% of mobile phase B. Elution was monitored at 210, 214, and 280 nm.

Antimicrobial activity analysis

The methods for small peptide antimicrobial activity were modified from (Yang et al., 2007). Fifty micro liters of *E. coli* culture in Luria-Bertani (LB) liquid medium with OD values between 0.4 and 0.5 were spread evenly on the pre-warmed (37 °C) LB solid medium. After the culture was dry (around one-half hour), twenty micro litters of 4 mM pure GAM1 (synthesized by Genscript) in PBS solution or double-distilled water (as negative control) was dropped on the LB medium. The plates were then incubated overnight at 37 °C.

Results and Discussion

Generation of the transgenic maize

Two proteins, the red fluorescent protein (RFP) and the maize endogenous seed storage protein 22-kDa α-zein, were selected for fusing to GAM1 in the present study. The constructs employed are shown in Figure 1. In the pYJ16 construct, GAM1 was fused to the C-terminus of RFP. RFP is a widely used reporter protein and the use of the RFP in maize cell biology has been developed (Shaner et al., 2004). Additionally, as a tag protein, there are various commercial purification kits available for purifying RFP tagged fusions. The pYJ15 construct carries a gene coding sequence for the maize native 22-kDa α-zein fused to the N-terminus of GAM1, driven by the 22-kDa α-zein promoter. The 22-kDa α-zein terminator was used as well. The α-zeins were the most abundant zein proteins, accounting for 75-85% of total zeins (Tatham et al., 1993). This α-zein fusion strategy was successfully used (D. Jackson, Cold Spring Harbor Laboratories). By fusing to the 22-kDa α-zein and using its original promoter and terminator, we expected GAM1 would achieve high expression via the native maize 22-kDa α-zein system.

Both constructs were transformed into maize HiII germplasm by using immature embryos via *Agrobacterium*-based methods. Eleven independent events each for A274 and A273 were recovered and brought to maturity in the greenhouse by crossing with pollen donors of an inbred maize line B73. Initial transgene expression screens were completed on both calli and T_1 seeds using PCR. The PCR and the PCR products sequencing results in the calli and T_1 seeds indicated the GAM1 DNA coding sequence was intact and the protein::GAM1 DNA coding sequence fusion was correctly made from transgenic lines carrying both constructs (data not shown).

Characterization of the RFP::GAM1 transgenic line A274

Visualization of the RFP in the RFP::GAM1 transgenic line. The fluorescence imaging screen was conducted on the transgenic maize A274 line. The RFP was successfully expressed in the A274 line. As shown in Figure 2, the positive segregates of RFP positive kernels are red under the fluorescent field, while the RFP negative segregate are not. The red fluorescence suggested the RFP in this RFP::GAM1 fusion protein was correctly folded. A total of 633 kernels from a total of 2,071 kernels (30%) screened from 11 events were red fluorescent (Table 1). Of 11 events, only two events—A274-9 and A274-11—showed 1:1 segregation as measured for RFP expression. These were also the lines in which red fluorescence was strongest. The remaining nine events had the RFP expression ranging from 15.8% (A274-6) to 38.7% (A274-8). We did not perform PCR analysis on the presence of transgenes. Therefore, it is unclear whether the lack of RFP expression was due to the lack of gene or gene silencing. Event A274-9 was chosen for further analyses described below, due to its strong RFP expression and abundance of seeds.

Western Blot of the RFP in the RFP::GAM1 transgenic line. To confirm the red fluorescence in seed, indeed, was due to the expression of the RFP, we screened seeds selected from two events (A274-3-2 and A274-6-3) by using Western Blot analysis. Eight randomly selected seeds from each event were individually ground powder. The total aqueous soluble (TAS) protein was extracted from the seed powder at a ratio of 10 μL of extraction buffer per mg of seed powder. Equal volume (20 μL) from each seed sample was analyzed in Coomassie Brilliant Blue staining or Western Blot. As seen in Figure 3A,

TAS protein from eight seeds for each event and 1 μg of the commercial RFP were shown by Coomassie staining. A band around 28 kDa, close to the anticipated RFP size (27.6 kDa), could be detected in three seeds from A274-3-2 and one seed from A274-6-3 by RFP antibody in the Western Blot (Figure 3A). One microgram of commercial RFP was loaded on to the gel as positive control. However, the smear RFP band is shown for all samples.

Subsequently, we used a RFP affinity column to concentrate the RFP fusion protein from the A274 seeds. A274-9-2, an event with stronger red fluoresce in seeds, was selected for RFP purification. Due to product inconsistency of the RFP affinity columns purchased from MBL International, only the first lot of the Anti-RFP magnetic beads worked. The purified protein was split into two equal aliquots and 15 μl each were loaded on to two SDS-PAGE gels. One gel was for Coomassie Brilliant Blue staining and the other was for the Western Blot. One microgram of the commercial RFP was also loaded onto the gel as a positive control. As shown in Figure 3B, a band around 28 kDa could be detected from A274-9-2 by the RFP antibody in the Western Blot. A population of proteins smaller than the 28 kDa could also be detected from the same lane (observed as a smear). It is possible these RFP antibody cross-reacting proteins were the RFP degradation products.

The molecular weight of the commercial RFP control was 27.6 kDa (cat # 4997-100 BioVision). The molecular weight of the fusion protein RFP::GAM1 was 33.3 kDa, predicted from the protein sequence (the molecular weight of the RFP without fusion with GAM1 is 28.9 kDa). There may be a slight difference in molecular weight between the A274-9-2 sample and the positive control in the Western Blot. However, the

difference was not obvious, due to the unfocused band and the resolution limit of the gel. Further investigation is needed to ascertain the size of the maize-derived RFP fusion protein.

In summary, the RFP Western Blot result indicated the red fluorescence in seed was correlated with the presence of RFP protein in seed protein extract, but the presence of GAM1 was undetermined.

The in-gel digestion/peptide mass fingerprinting of the RFP::GAM1 fusion protein. To determine if GAM1 was also expressed with the fusion protein RFP, the band of A274-9-2 derived RFP in the Coomassie Brilliant Blue stained SDS-PAGE gel was cut out and collected. The gel piece was then digested with trypsin. After digestion, the peptides from the sample were released from the gel and analyzed by tandem mass spectrometry (MS/MS). The peptide sequence was identified by its pattern from the MS results. Thus, the target protein was determined by the sequenced peptides via a protein database search. The Protein Facility at Iowa State University completed this in-gel digestion/peptide mass fingerprinting method. By using this method, we expected to identify the GAM1 sequence.

The GAM1 however was undetected on several attempts. The peptide sequences identified from the MS/MS are always from a common contamination source, Keratin. We speculate that the desired peptide sequences were not detected in our samples due to a low concentration of the fusion protein. In an attempt to achieve higher concentration and better purity of the RFP fusion protein from the seed extracts, we tested a number of RFP purification products, such as Protein A MagBeads (Cat # L00273, GenScript), RFP-tag Antibody (pAb, Rabbit Cat # A00682-40, GenScript), in addition to the

purification kit from MBL International Company described above. Nevertheless, we could not obtain a better yield for RFP fusion protein. After discussions with Dr. Charles E. Glatz (Department of Chemical and Biological Engineering, Iowa State University), we decided to use HPLC as an alternative method to detect GAM1.

The HPLC of the RFP::GAM1 fusion protein. Dr. Glatz kindly allowed us use of the HPLC in his lab to detect the RFP::GAM1, since the *E. coli* derived GAM1 has been successfully identified by using HPLC. Two types of protein samples were used for this analysis. One sample type was protein purified by the RFP affinity column from A274-9-2. The other sample was the same purified sample, but treated with TEV protease. By using the TEV treatment, the GAM1 was expected to cleave from the fusion protein and be detected by HPLC. Because of the limited amount of protein, we did not measure the concentration prior to sample loading. Fifty microliters of each protein sample (treated with or without TEV) was subjected to HPLC analysis. We also included a synthetic GAM1 (synthesized by GenScript) as a positive control.

Only the 1 μg synthesized GAM1 was detected from the HPLC analysis (Figure 4). Neither the RFP::GAM1 fusion protein nor the GAM1 peptide (from TEV-treated protein sample) could be detected. It seemed the expression of the fusion protein was below the detection limit.

Further research is needed to verify the expression of GAM1 in the A274 line. Not all RFP::GAM1 transgenic events were analyzed in this study. Only the A274-9-2 was characterized because it had the most seeds with strong red fluorescence. More or even all events should be analyzed. Other commercially-available RFP purification kits should also be utilized. Some advanced instruments with higher detection capability, such

as High-Resolution Mass Spectrometry (HRMS), should be considered ($1,000 per sample) as well.

Characterization of the α-zein::GAM1 transgenic line A273

In this fusion protein strategy, GAM1 was fused to the C-terminus of maize native 22-kDa α-zein, under the control of the 22-kDa α-zein promoter and terminator. By fusing to the 22-kDa α-zein and using its original promoter and terminator, GAM1 was expected to achieve high expression by using the machinery of the most abundant zein proteins.

PCR-positive seeds from two events, A273-2-3 and A273-4-2, were selected for characterization because they had large, healthy looking seeds. There is no commercial purification kit or antibody available for 22-kDa α-zein. Dr. Paul Scott, USDA and Department of Agronomy, Iowa State University, kindly offered us an antibody for zeins. The Western Blot analysis was achieved by using this antibody, but multiple bands were shown and no difference could be determined between the transgenic positive and negative samples (Figure 5). We also checked to determine if the 22-kDa α-zein might switch from alcohol-soluble to aqueous-soluble, due to fusion with GAM1. Therefore, the TAS protein extracted from the A273 transgenic positive sample was also analyzed by Western Blot, however, no zein was detected.

With help from the Protein Facility at Iowa State University, the band between 25 to 37 kDa (the predicted 22-kDa α-zein::GAM1 fusion molecular weight is 30.6 kDa) from the Coomassie Brilliant Blue stained SDS-PAGE gel were collected and analyzed

using the in-gel digestion/peptide mass fingerprinting method. Nevertheless, GAM1 could not be detected.

HPLC was conducted to analyze the 22-kDa α-zein fused GAM1 in the A273-2-3 line. The total zein proteins were extracted from PCR positive A273-2-3 mature seeds and analyzed by HPLC (Figure 6). If the α-zein::GAM1 was successfully produced in detectable quantity, one additional peak was expected in the A273-2-3 sample when compared to its negative segregants. However, no difference was determined between the PCR positive and negative samples by comparison of the HPLC results. To perform more detailed analysis, each peak from HPLC was collected and further analyzed individually by using the mass spectrometer (MS) in the Protein Facility at Iowa State University. Nevertheless, the expected fusion protein still could not be identified from a high background, possiblly due to a limitation of the MS (Figure 6B). All 22 peaks were analyzed and an example of a peak at 37 min is shown in Figure 6B.

So far, all methods attempted failed to verify the expression of the α-zein::GAM1 fusion protein in the A273 line. The α-zein fusion strategy might be reconsidered. It has been reported that a 30-amino-acid long peptide hormone GLP-1 was successfully expressed in rice when fused to the C-terminus of GFP (green fluorescent protein) (Yasuda et al., 2005). The fusion protein was under the control of the rice storage protein glutelin GluB-1 promoter, and the glutelin GluB-1 terminator was also used. The signal peptide derived from glutelin GluB-1 and the ER retention sequence KDEL were included in the construct to lead the fusion protein and accumulate in the ER. The accumulation of GFP::GLP-1was detected at the RNA level by Northern Blot in immature seeds and at the protein level by Western Blot (anti GLP-1 or anti-GFP) in

mature seeds (Yasuda et al., 2005). Another successful example was reported by expressing a 39-amino-acid-long antimicrobial peptide sarcotoxin IA (ST) in tobacco by fused to the β-glucuronidase (GUS) (Okamoto et al., 1998). The high expression of the fusion protein was achieved by either a ST::GUS or GUS::ST fusion strategy and with or without a signal peptide of *PRla* gene of *Nicotiana tabacum*. All constructs were under the control of the 35S promoter and the nos terminator. The expression of the fusion protein was analyzed by both GUS assay and anti-ST Western Blot (Okamoto et al., 1998). These studies suggested that either using an endogenous gene promoter or a strong constitutive promoter leads to high expression of fusion proteins. Adding a signal peptide and KDEL is also a good strategy. The C- or N- terminus fusion did affect the expression. In addition, using an exogenous fusion to partner with the effective detection methods available (e.g. GFP or GUS) may facilitate identification and purification.

In theory, the two strategies designed in the present study should work. However, using an endogenous fusion partner (i.e. α-zein) may not be effective because the endogenous zein protein was abundant and no effective method exists for detecting or purifying either α-zein or GAM1. Future work should verify the RNA expression of the α-zein::GAM1 by using Northern Blot or RT-PCR in immature A773 seeds.

Antimicrobial function

Peptides with antimicrobial properties are found in all kingdoms of life as part of the defense system against microbial invasion (Uzarski and Mello, 2012). Many recombinant antimicrobial peptides also have been successfully produced from various heterologous hosts (Ingham and Moore, 2007). The synthesized GAM1 (by GenScript USA, Inc.) was tested for its antimicrobial ability, according to the methods previously

described in (Yang et al., 2007). In short, 20 μL of 4 mM pure GAM1 in PBS solution was dropped on the center of a LB plate with *E. coli* spread evenly. If GAM1 had antimicrobial activity, the area GAM1 covered on the medium showed less or no bacteria colonies after overnight incubation at 37^{o}C. No bacterial growth inhibition was observed on the GAM1 plates when compared with water control plates. This preliminary experiment suggested the GAM1 peptide had no antimicrobial function against laboratory strain *E. coli* DH5α used for recombinant DNA work. Future experiments including other gram positive or negative bacterial strains or fungi should be conducted.

Summary

GAM1 was reported through production in recombinant *E. coli* fused with the maltose binding protein (MBP) tag (Kaar et al., 2009). The bioactivity test of the *E. coli* expressed GMA1 indicated this biosurfactant peptide was surface active and able to induce switching between foam-stabilizing and foam-destabilizing states (Kaar et al., 2009). Due to the contamination risk and low product quality of recombinant bacteria expression system, the transgenic plants have been considered as alternative production systems. Here, we report on an attempt to express GAM1 in maize seeds by using a fusion protein strategy. The ultimate goal was to recover the fusion proteins from maize seeds and utilize the seed remnant for fuel ethanol production.

Two fusion partners, the RFP and 22-kDa α-zein, were utilized to facilitate purification and enhance expression of GMA1. Transgenic maize lines carrying RFP::GAM1 and α-zein::GAM1 were successfully generated, and the correct fusion sequences were confirmed by DNA sequence analysis and PCR. Seeds in RFP::GAM1

transgenic maize lines (A274) displaying red fluorescence were positive for RFP in Western Blot analysis, suggesting the RFP in these seeds were produced and correctly folded. However, we could not verify the presence of the GAM1 in both RFP::GAM1 and Zein::GAM1 fusion lines, despite efforts to use various analytical methods, including in-gel digestion/peptide mass fingerprinting, HPLC, and MS.

It is unclear why we were not able to detect GAM1 in both fusion proteins. First, we confirmed the fusion sequences were correctly assembled in the vector and intact transgenes were present in transgenic seeds. However, we did not check the transgene transcripts in seeds. We need to perform reverse transcription PCR (RT-PCR) to determine if the entire fusion transcripts can be detected in seeds, even though seeds possess a low quantity of RNA. Second, we detected red fluorescence in seeds carrying RFP::GAM1 fusion protein. We know at least the RFP was produced and correctly folded. Because we could not detect the presence of GAM1, it was possible the fusion protein did not become fully processed in the plant seeds due to unknown reasons; only the N-terminal part of fusion protein, RFP, was produced. It was also possible the overall low expression level of RFP::GAM1 fusion protein was below the detection levels of the methods used.

Future work should include RT-PCR screening on all transgenic maize seeds to determine whether the full length fusion transcripts are present in seed. If they can be detected, one can proceed with protein/peptide detection by using an instrument with higher detection capabilities. If no full-length fusion transcripts are detected, one will need to redesign the production strategy for this small peptide. Maize seeds may not be the best system for small peptide production.

Acknowledgments

We thank Dr. Charles Glatz, Dr. Paul Scott, William Colonna, and Dr. Keunsub Lee for helpful discussions, and sharing materials and equipment for the HPLC and Western blot. We thank Margie Carter, Siquan Luo, and Joel Nott in the Protein Facility, and Ann Perera in W.M. Keck Metabolomics Research Laboratory at Iowa State University for helpful discussions and suggestions. We also thank Angela Nguyen for her technical assistance. This project was supported by USDA Special grant IOW05228.

References

Achmuller, C., Kaar, W., Ahrer, K., Wechner, P., Hahn, R., Werther, F., Schmidinger, H., Cserjan-Puschmann, M., Clementschitsch, F., Striedner, G., Bayer, K., Jungbauer, A. and Auer, B. (2007) N-pro fusion technology to produce proteins with authentic N termini in E-coli. *Nature Methods* **4**, 1037-1043.

Anzai, H., Yoneyama, K. and Yamaguchi, I. (1989) Transgenic tobacco resistant to a bacterial disease by the detoxification of a pathogenic toxin. *Molecular & General Genetics* **219**, 492-494.

Azzoni, A., Kusnadi, A., Miranda, E. and Nikolov, Z. (2002) Recombinant aprotinin produced transgenic corn seed: Extraction and purification studies. *Biotechnology and Bioengineering* **80**, 268-276.

Bailey, M., Woodard, S., Callaway, E., Beifuss, K., Magallanes-Lundback, M., Lane, J., Horn, M., Mallubhotla, H., Delaney, D., Ward, M., Van Gastel, F., Howard, J.

and Hood, E. (2004) Improved recovery of active recombinant laccase from maize seed. *Applied Microbiology and Biotechnology* **63**, 390-397.

Banat, I., Franzetti, A., Gandolfi, I., Bestetti, G., Martinotti, M., Fracchia, L., Smyth, T. and Marchant, R. (2010) Microbial biosurfactants production, applications and future potential. *Applied Microbiology and Biotechnology* **87**, 427-444.

Banat, I., Makkar, R. and Cameotra, S. (2000) Potential commercial applications of microbial surfactants. *Applied Microbiology and Biotechnology* **53**, 495-508.

Clapes, P. and Infante, M. (2002) Amino acid-based surfactants: Enzymatic synthesis, properties and potential applications. *Biocatalysis and Biotransformation* **20**, 215-233.

Deleu, M., Razafindralambo, H., Popineau, Y., Jacques, P., Thonart, P. and Paquot, M. (1999) Interfacial and emulsifying properties of lipopeptides from Bacillus subtilis. *Colloids and Surfaces a-Physicochemical and Engineering Aspects* **152**, 3-10.

Dexter, A., Malcolm, A. and Middelberg, A. (2006) Reversible active switching of the mechanical properties of a peptide film at a fluid-fluid interface. *Nature Materials* **5**, 502-506.

Dexter, A. and Middelberg, A. (2007) Switchable peptide surfactants with designed metal binding capacity. *Journal of Physical Chemistry C* **111**, 10484-10492.

Dexter, A. and Middelberg, A. (2008) Peptides as functional surfactants. *Industrial & Engineering Chemistry Research* **47**, 6391-6398.

Dougherty, W., Carrington, J., Cary, S. and Parks, T. (1988) Biochemical and mutational analysis of a plant-virus polyprotein cleavage site. *Embo Journal* **7**, 1281-1287.

Flint-Garcia, S., Bodnar, A. and Scott, M. (2009) Wide variability in kernel composition, seed characteristics, and zein profiles among diverse maize inbreds, landraces, and teosinte. *Theoretical and Applied Genetics* **119**, 1129-1142.

Frame, B.R., Shou, H.X., Chikwamba, R.K., Zhang, Z.Y., Xiang, C.B., Fonger, T.M., Pegg, S.E.K., Li, B.C., Nettleton, D.S., Pei, D.Q. and Wang, K. (2002) Agrobacterium tumefaciens-mediated transformation of maize embryos using a standard binary vector system. *Plant Physiology* **129**, 13-22.

Francois, I., De Bolle, M., Dwyer, G., Goderis, I., Woutors, P., Verhaert, P., Proost, P., Schaaper, W., Cammue, B. and Broekaert, W. (2002) Transgenic expression in Arabidopsis of a polyprotein construct leading to production of two different antimicrobial proteins. *Plant Physiology* **128**, 1346-1358.

Gomord, V., Denmat, L., FitchetteLaine, A., SatiatJeunemaitre, B., Hawes, C. and Faye, L. (1997) The C-terminal HDEL sequence is sufficient for retention of secretory proteins in the endoplasmic reticulum (ER) but promotes vacuolar targeting of proteins that escape the ER. *Plant Journal* **11**, 313-325.

Gordonkamm, W., Spencer, T., Mangano, M., Adams, T., Daines, R., Start, W., Obrien, J., Chambers, S., Adams, W., Willetts, N., Rice, T., Mackey, C., Krueger, R., Kausch, A. and Lemaux, P. (1990) Transformation of maize cells and regeneration of fertile transgenic plants. *Plant Cell* **2**, 603-618.

Grangemard, I., Peypoux, F., Wallach, J., Das, B., Labbe, H., Caille, A., Genest, M., Maget-Dana, R., Ptak, M. and Bonmatin, J. (1997) Lipopeptides with improved properties: Structure by NMR, purification by HPLC and structure-activity

relationships of new isoleucyl-rich surfactins. *Journal of Peptide Science* **3**, 145-154.

Hood, E., Helmer, G., Fraley, R. and Chilton, M. (1986) The hypervirulence of agrobacterium-tumefaciens a281 is encoded in a region of ptibo542 outside of transfer DNA. *Journal of Bacteriology* **168**, 1291-1301.

Ingham, A. and Moore, R. (2007) Recombinant production of antimicrobial peptides in heterologous microbial systems. *Biotechnology and Applied Biochemistry* **47**, 1-9.

Itakura, K., Hirose, T., Crea, R., Riggs, A., Heyneker, H., Bolivar, F. and Boyer, H. (1977) Expression in escherichia-coli of a chemically synthesized gene for hormone somatostatin. *Science* **198**, 1056-1063.

Johnson, L.A., May, J.B. and White, P.J. (2003) Wet milling: The basis for corn biorefineries. In: *Corn: Chemistry and Technology.* American Association of Cereal Chemists, 449-494.

Kaar, W., Hartmann, B., Fan, Y., Zeng, B., Lua, L., Dexter, A., Falconer, R. and Middelberg, A. (2009) Microbial Bio-Production of a Recombinant Stimuli-Responsive Biosurfactant. *Biotechnology and Bioengineering* **102**, 176-187.

Kanodia, S. and Kast, W. (2008) Peptide-based vaccines for cancer: realizing their potential. *Expert Review of Vaccines* **7**, 1533-1545.

Kapust, R. and Waugh, D. (1999) Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science* **8**, 1668-1674.

Kichler, A., Leborgne, C., Marz, J., Danos, O. and Bechinger, B. (2003) Histidine-rich amphipathic peptide antibiotics promote efficient delivery of DNA into

mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 1564-1568.

Ladurner, A. and Fersht, A. (1997) Glutamine, alanine or glycine repeats inserted into the loop of a protein have minimal effects on stability and folding rates. *Journal of Molecular Biology* **273**, 330-337.

Lee, E., Kim, H. and Lee, E. (2005) Recombinant biocatalytic and cell-free synthesis of HIV fusion inhibitor. *Journal of Industrial and Engineering Chemistry* **11**, 515-521.

Li, Y. (2009) Carrier proteins for fusion expression of antimicrobial peptides in Escherichia coli. *Biotechnology and Applied Biochemistry* **54**, 1-9.

Ma, J., Drake, P. and Christou, P. (2003) The production of recombinant pharmaceutical proteins in plants. *Nature Reviews Genetics* **4**, 794-805.

Malcolm, A., Dexter, A. and Middelberg, A. (2006) Foaming properties of a peptide designed to form stimuli-responsive interfacial films. *Soft Matter* **2**, 1057-1066.

Menkhaus, T., Bai, Y., Zhang, C., Nikolov, Z. and Glatz, C. (2004) Considerations for the recovery of recombinant proteins from plants. *Biotechnology Progress* **20**, 1001-1014.

Mine, Y. and Shahidi, F. (2006) Nutraceutical proteins and peptides in health and disease. In: *Nutraceutical Science and Technology 4* p. 668 p. Boca Raton: CRC/Taylor and Francis.

Morreale, G., Lee, E., Jones, D. and Middelberg, A. (2004) Bioprocess-centered molecular design (BMD) for the efficient production of an interfacially active peptide. *Biotechnology and Bioengineering* **87**, 912-923.

Murray, M.G. and Thompson, W.F. (1980) Rapid Isolation of High Molecular-Weight Plant DNA. *Nucleic Acids Research* **8**, 4321-4325.

Nitschke, M. and Costa, S. (2007) Biosurfactants in food industry. *Trends in Food Science & Technology* **18**, 252-259.

Okamoto, M., Mitsuhara, I., Ohshima, M., Natori, S. and Ohashi, Y. (1998) Enhanced expression of an antimicrobial peptide sarcotoxin IA by GUS fusion in transgenic tobacco plants. *Plant and Cell Physiology* **39**, 57-63.

Paraman, I., Moeller, L., Scott, M., Wang, K., Glatz, C. and Johnson, L. (2010) Utilizing Protein-Lean Coproducts from Corn Containing Recombinant Pharmaceutical Proteins for Ethanol Production. *Journal of Agricultural and Food Chemistry* **58**, 10419-10425.

Paz, M., Shou, H., Guo, Z., Zhang, Z., Banerjee, A. and Wang, K. (2004) Assessment of conditions affecting Agrobacterium-mediated soybean transformation using the cotyledonary node explant. *Euphytica* **136**, 167-179.

Peypoux, F., Bonmatin, J. and Wallach, J. (1999) Recent trends in the biochemistry of surfactin. *Applied Microbiology and Biotechnology* **51**, 553-563.

Piers, K., Brown, M. and Hancock, R. (1993) Recombinant-dna procedures for producing small antimicrobial cationic peptides in bacteria. *Gene* **134**, 7-13.

Ramessar, K., Sabalza, M., Capell, T. and Christou, P. (2008) Maize plants: An ideal production platform for effective and safe molecular pharming. *Plant Science* **174**, 409-419.

Sekhon, B. (2006) Biosurfactants: An overview. *National Academy Science Letters-India* **29**, 317-332.

Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N.G., Palmer, A.E. and Tsien, R.Y. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology* **22**, 1567-1572.

Singh, A., Van Hamme, J. and Ward, O. (2007) Surfactants in microbiology and biotechnology: Part 2. Application aspects. *Biotechnology Advances* **25**, 99-121.

Soave, C., Tardani, L., Difonzo, N. and Salamini, F. (1981) Zein level in maize endosperm depends on a protein under control of the opaque-2 and opaque-6 loci. *Cell* **27**, 403-410.

Tatham, A., Field, J., Morris, V., Ianson, K., Cardle, L., Dufton, M. and Shewry, P. (1993) Solution conformational-analysis of the alpha-zein proteins of maize. *Journal of Biological Chemistry* **268**, 26253-26259.

Uchimiya, H., Iwata, M., Nojiri, C., Samarajeewa, P., Takamatsu, S., Ooba, S., Anzai, H., Christensen, A., Quail, P. and Toki, S. (1993) Bialaphos treatment of transgenic rice plants expressing a bar gene prevents infection by the sheath blight pathogen (rhizoctonia-solani). *Bio-Technology* **11**, 835-836.

Uzarski, J. and Mello, C. (2012) Detection and Classification of Related Lipopolysaccharides via a Small Array of Immobilized Antimicrobial Peptides. *Analytical Chemistry* **84**, 7359-7366.

Van Hamme, J., Singh, A. and Ward, O. (2006) Physiological aspects - Part 1 in a series of papers devoted to surfactants in microbiology and biotechnology. *Biotechnology Advances* **24**, 604-620.

Vogt, T. and Bechinger, B. (1999) The interactions of histidine-containing amphipathic helical peptide antibiotics with lipid bilayers - The effects of charges and pH. *Journal of Biological Chemistry* **274**, 29115-29121.

Woodard, S., Mayor, J., Bailey, M., Barker, D., Love, R., Lane, J., Delaney, D., McComas-Wagner, J., Mallubhotla, H., Hood, E., Dangott, L., Tichy, S. and Howard, J. (2003) Maize (Zea mays)-derived bovine trypsin: characterization of the first large-scale, commercial protein product from transgenic plants. *Biotechnology and Applied Biochemistry* **38**, 123-130.

Xu, X., Gan, Q., Clough, R., Pappu, K., Howard, J., Baez, J. and Wang, K. (2011) Hydroxylation of recombinant human collagen type I alpha 1 in transgenic maize co-expressed with a recombinant human prolyl 4-hydroxylase. *Bmc Biotechnology* **11**, 69.

Xu, Z., Zhong, Z., Huang, L., Peng, L., Wang, F. and Cen, P. (2006) High-level production of bioactive human beta-defensin-4 in Escherichia coli by soluble fusion expression. *Applied Microbiology and Biotechnology* **72**, 471-479.

Yang, X., Xiao, Y., Wang, X. and Pei, Y. (2007) Expression of a novel small antimicrobial protein from the seeds of motherwort (Leonurus japonicus) confers disease resistance in tobacco. *Applied and Environmental Microbiology* **73**, 939-946.

Yasuda, H., Tada, Y., Hayashi, Y., Jomori, T. and Takaiwa, F. (2005) Expression of the small peptide GLP-1 in transgenic plants. *Transgenic Research* **14**, 677-684.

Zhou, Q., Luo, X., Ye, L. and Xi, T. (2007) High-level production of a novel antimicrobial peptide perinerin in Escherichia coli by fusion expression. *Current Microbiology* **54**, 366-370.

Table 1 Generation of transgenic maize line A274 (RFP::GMA1)

	PTF ID	Event	Plant	# total seed	*# red fluorescent seed	% red fluorescent	**Fluorescent intensity
	A274	1	2	16	2	12.5%	
1	A274	1	3	47	14	29.8%	
		subtotal		63	16	25.4%	
	A274	3	1	109	30	27.5%	
	A274	3	2	126	44	34.9%	strong
2	A274	3	3	1	1	100.0%	
	A274	3	4	174	53	30.5%	strong
		subtotal		410	128	31.2%	
	A274	5	1	21	10	47.6%	
3	A274	5	2	32	9	28.1%	
	A274	5	3	2	0	0.0%	
		subtotal		55	19	34.5%	
	A274	6	1	109	0	0.0%	
	A274	6	2	125	28	22.4%	
4	A274	6	3	117	25	21.4%	
	A274	6	4	3	3	100.0%	
		subtotal		354	56	15.8%	
	A274	7	1	113	18	15.9%	
	A274	7	2	2	2	100.0%	
5	A274	7	3	23	2	8.7%	
	A274	7	4	36	9	25.0%	
		subtotal		174	31	17.8%	
	A274	8	1	20	5	25.0%	
	A274	8	2	1	0	0.0%	
6	A274	8	3	3	2	66.7%	
	A274	8	4	7	5	71.4%	
		subtotal		31	12	38.7%	
	A274	9	1	116	44	37.9%	strong
	A274	9	2	67	51	76.1%	very strong
7	A274	9	3	82	47	57.3%	strong
	A274	9	4	116	49	42.2%	strong
		subtotal		381	191	50.1%	
	A274	10	2	31	8	25.8%	
8	A274	10	3	53	15	28.3%	
		subtotal		84	23	27.4%	
	A274	11	1	45	20	44.4%	very strong
	A274	11	2	29	13	44.8%	very strong
9	A274	11	3	26	14	53.8%	very strong
	A274	11	4	33	13	39.4%	very strong
		subtotal		133	60	45.1%	
	A274	12	1	86	21	24.4%	strong
	A274	12	2	78	30	38.5%	
10	A274	12	3	89	17	19.1%	
	A274	12	4	97	29	29.9%	strong
		subtotal		350	97	27.7%	
11	A274	13	1	36	0	0.0%	
		total		2071	633	30.6%	

* determined by using a stereoscope coupled with RFP filter

** intensity roughly determined by naked eyes

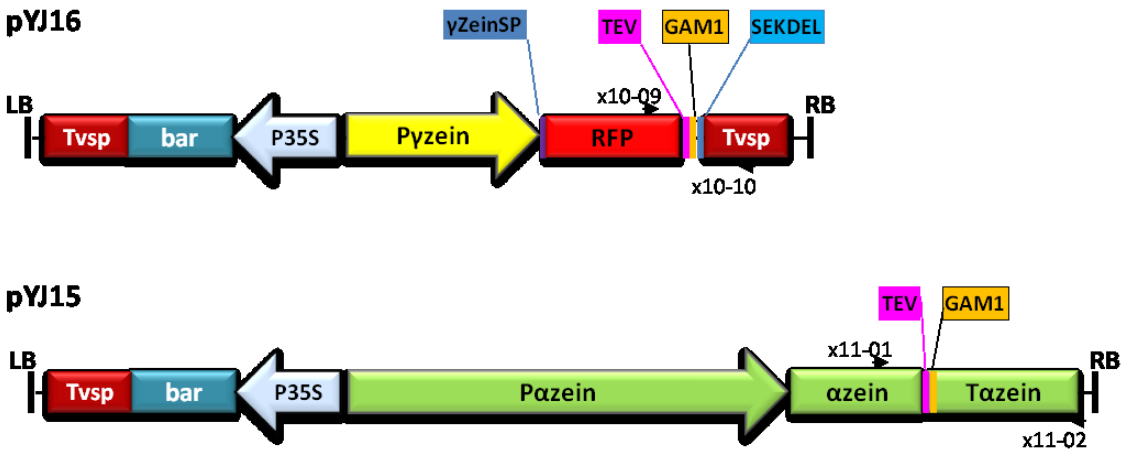

Figure 1. Constructs used in this study. pYJ16, construct for generating transgenic maize line A274. Pγzein, 27-kDa γ-zein promoter; γzeinSP, 27-kDa γ-zein signal peptide; RFP, the red fluorescent protein; TEV, tobacco etch virus protease recognition site; GAM1, the biosurfactant peptide; SEKDEL, endoplasmic reticulum retention sequence; x10-09 and x10-10, primers for verifying the presence of RFP::GAM1 in pYJ16; pYJ15, construct for generating transgenic maize line A273. Pαzein, 22-kDa α-zein promoter; αzein, 22-kDa α-zein gene coding region; Tαzein, 22-kDa α-zein gene terminator; x11-01 and x11-02, primers for verifying the presence of α-zein::GAM1 in pYJ15; Tvsp, soybean vegetative storage protein gene terminator; P35S, 35S promoter; bar, bar gene coding region; LB, left border of *Agrobacterium* T-DNA; RB, right border of *Agrobacterium* T-DNA

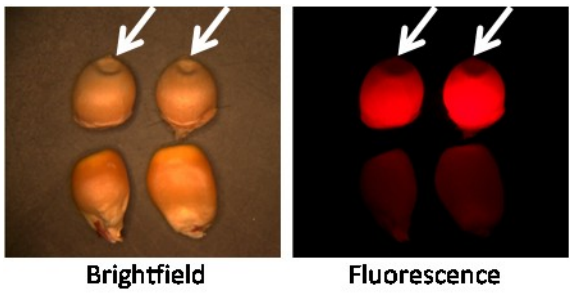

Figure 2. Bright field and fluorescent imaging of A274-9-2 (RFP::GAM1) kernels. White arrows indicate the segregate of two RFP positive kernels. The other two are the segregate of RFP negative kernels

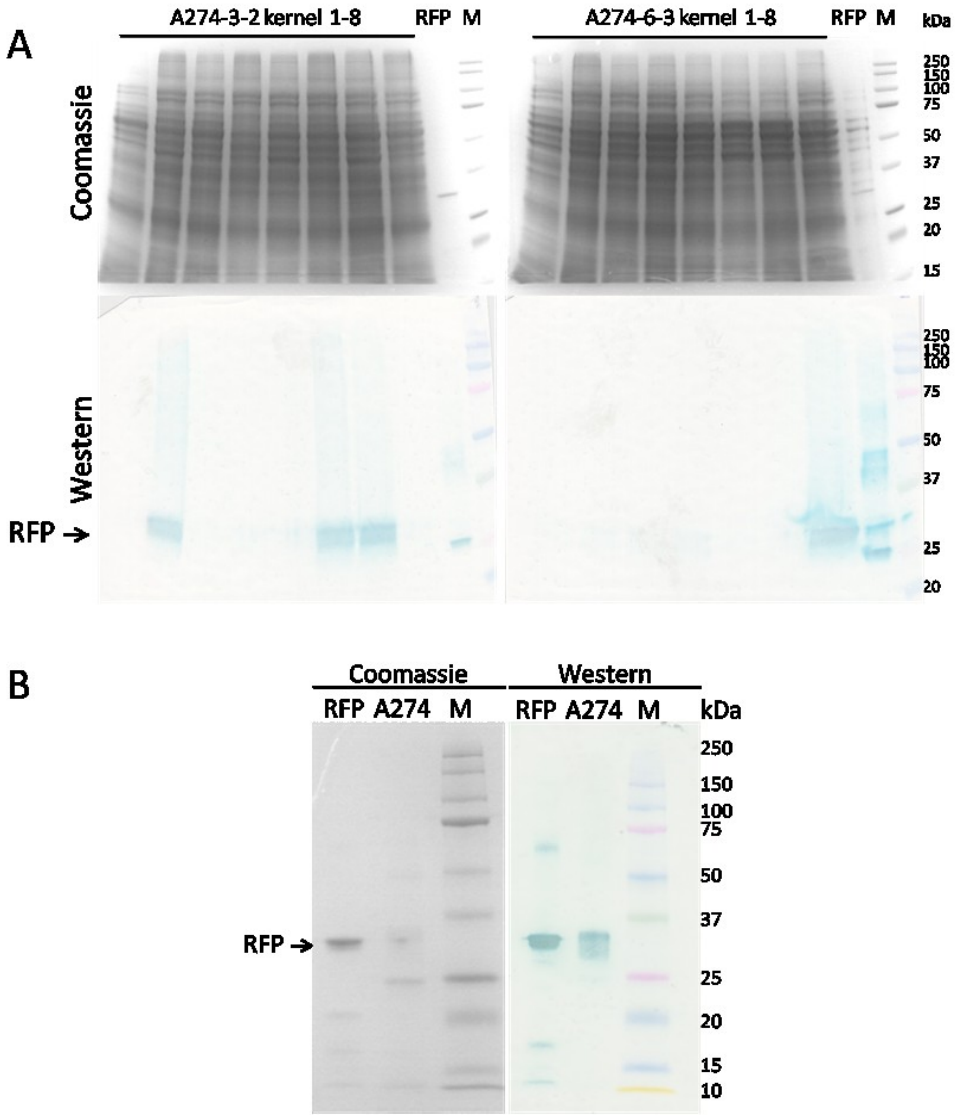

Figure 3. Coomassie Brilliant Blue staining and Western blot of A274 (RFP::GAM1) line. (A) Coomassie Brilliant Blue staining and Western blot of total aqueous soluble protein extracted from seed powders of eight kernels from each —274-3-2 and A274-6-3. (B) Coomassie Brilliant Blue staining and Western blot of RFP affinity column purified total aqueous soluble protein from A274-9-2 (denoted as A274 on the top of both Coomassie and Western lanes). RFP, 1 µg of the commercial RFP control; M, the molecular weight standard. Anti-RFP antibody was used to detect the RFP in Western blot.

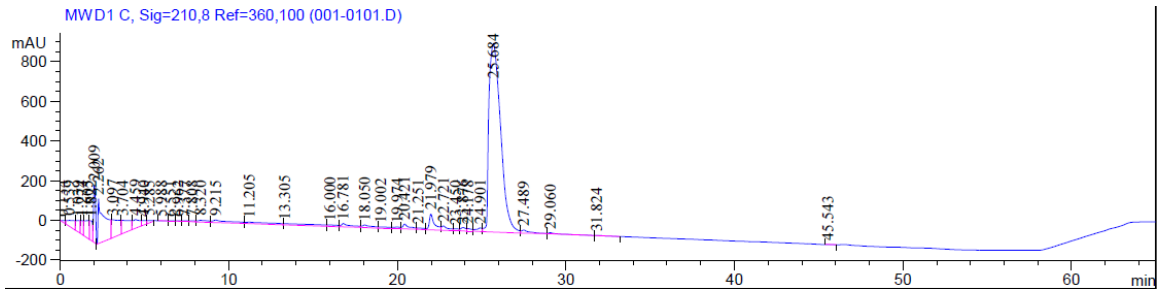

Figure 4. GAM1 HPLC. 1 µg synthesized GAM1 was detected by HPLC around 28 min. The wavelength was 210 nm.

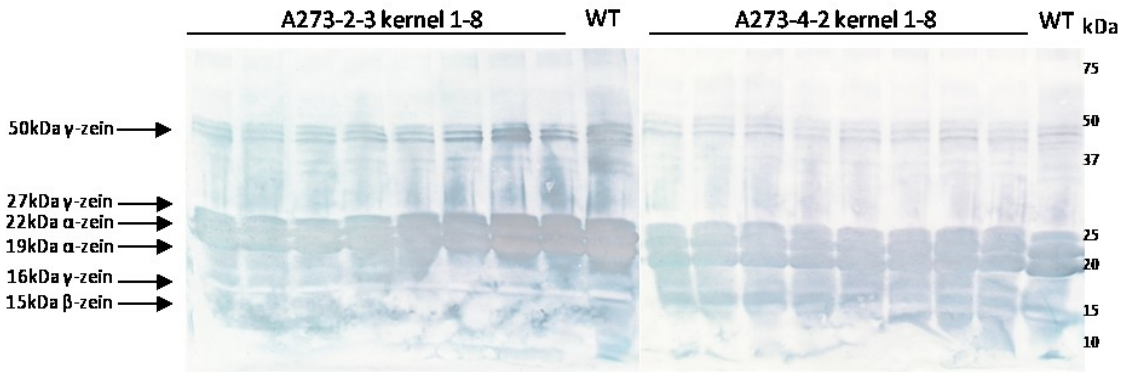

Figure 5. Western Blot of α-zein::GAM1. Anti-zein rabbit polyclonal antibody (from Dr. Paul Scott) is used to detect zeins from A273-2-3 and A273-4-2. WT, wild type control B73 line.

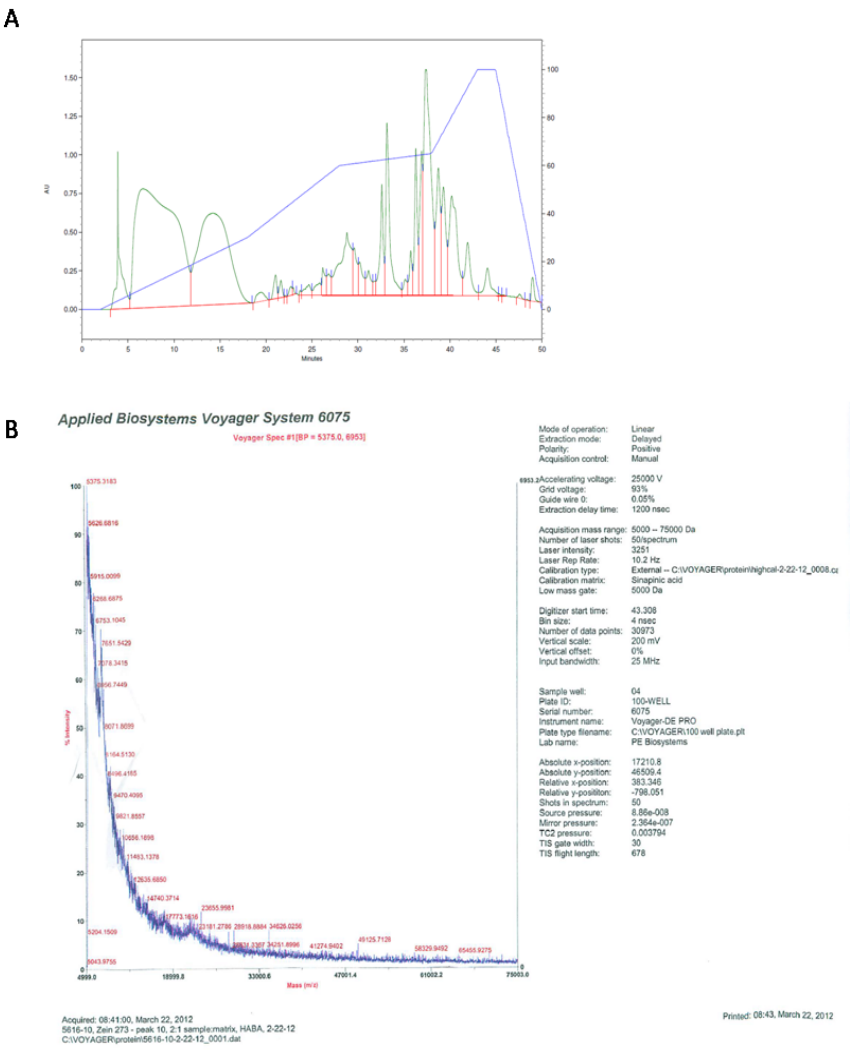

Figure 6. Transgenic maize line A273-2-3 zein Liquid Chromatography Mass Spectrometry (LCMS). Total zein proteins were extracted from A273-2-3 PCR-positive seed and analyzed by LCMS. (A) Zeins separated by the C18 column (5 microns, 2.1 mm X 250 mm). The wavelength for detecting zeins is 214 nm. (B) One MS result of the peak appears around 37 min from liquid chromatography. No significant protein was identified from the background in this peak. All other individual peaks were also analyzed by MS, but not shown here, due to the length of the figure. No significant proteins were identified from all peaks.

CHAPTER 5: STRATEGIES FOR THE PRODUCTION OF MAIZE-DERIVED PHARMACEUTICALS USING CYTOPLASM MALE STERILE LINES: *IN VITRO* TISSUE CULTURE/TRANSFORMATION AND FIELD BREEDING APPROACHES

Modified from a paper published in Maydica 2009, 54:199-210

Kan Wang, Bronwyn Frame, Xing Xu, Lorena Moeller, Kendall Lamkey, Roger Wise

Abstract

Plant-made pharmaceuticals (PMPs) offer promise as efficient and cost-effective products for the treatment of human and animal diseases. An advantage of producing pharmaceuticals in maize is the large storage capacity and stability for proteins and starch in seed, allowing for manufacturing recombinant proteins such as antigens and antibodies. Other advantages of the maize system include safety, high yields, and scalability of production and processing. However, the benefits of this technology must be balanced against potential health and environmental risks that may be associated with its use. Because PMPs presently have no provision for regulatory tolerance, their inadvertent occurrence in foods and feeds remains an important economic consideration, even when the health and environmental risks are low. Pollen drift is considered a source of potential contamination of maize-made pharmaceuticals in the food chain. In addition to physical and temporal isolation requirements, open field pharmaceutical maize production also calls for controlled

pollen release. Here, we describe two strategies to address the issue of transgenic pollen drift. First, we describe the development and genetic transformation of a tissue culture-amenable male-sterile line using biolistic- or *Agrobacterium*-mediated transformation methods. Secondly, we describe the introgression of a transgene from male-fertile transgenic maize to male-sterile germplasm by conventional breeding. After six seasons of breeding, this second strategy allows us to obtain 100% transgenic seeds from an open-field production using a non-transgenic line as the pollinator.

Introduction

Maize grain is used as feedstock for many large volume industrial products (e.g. ethanol, biodiesel, poly-lactic acid, sweeteners) and has been demonstrated to be an effective expression system for functional proteins of prokaryotic (Bailey et al., 2004; Chikwamba et al., 2002a; Streatfield et al., 2001; Witcher et al., 1998;); viral (Streatfield et al., 2001) and eukaryotic (Hood et al., 1997; Yang et al., 2002; Zhong et al., 1999) origins. Previously we have successfully produced a transgenic maize line expressing a B-subunit protein of *Escherichia coli* heat labile enterotoxin (LT-B) (Chikwamba et al., 2002a). Mice fed with maize-derived LT-B were protected against diarrhea-inducing *E coli* toxins and showed reduced symptoms compared to control mice fed with non-transgenic maize meal (Chikwamba et al., 2002a).

There are several benefits of producing antigens for use as edible vaccines in maize (Streatfield, 2007; Ramessar et al., 2008). Maize is a major crop world wide used for food, feed and fuel. As food and feed, maize is well tolerated by both humans and animals since it is non-toxic and non-allergenic. Compared to bacterial and mammalian systems, the use of

maize for producing recombinant proteins greatly reduces the possibility of contamination with mammalian pathogens. Maize yields are high and its seed is a natural protein storage site that can be harnessed as an efficient protein production factory. According to Hood and coworkers (Hood et al., 1999), maize can be used to produce foreign proteins at rates that are in excess of 2 kilograms per acre at a cost of a few cents per milligram. Breeding techniques can be used in this crop to enhance foreign protein expression (Chikwamba et al., 2002b; Hood et al., 1997; Streatfield et al., 2001; Witcher et al., 1998). Moreover, the infrastructure for large scale grain production and seed processing are well established for the crop in the United States.

Since maize is one of the most important crops globally, there are many issues that arise regarding its use as a source of biopharmaceuticals. Major concerns, especially in the case of vaccine-producing maize plants, are whether the antigens will inadvertently enter the food chain and what measures are in place to prevent this (Ripplinger et al., 2009). There are three broad categories of grain handling systems used in the U.S.: commodity, identity-preserved, and seed (Strayer, 2004). These systems vary in their purity standards, which are primarily based on end-user requirements. The commodity system is used when there are no requirements for differentiation by end users and the identity-preserved system is used when end-users require specific grain characteristics. The seed system is used specifically for producing the seed to be planted for commodity and identity-preservation systems (Strayer, 2002; 2004).

One major concern with regard to cross-contamination in these systems is the potential for transgenic maize pollen to drift to surrounding non-transgenic maize fields. Several measures can be taken to prevent pollen contamination from a vaccine-producing

maize plant. First, any APHIS (Animal and Plant Health Inspection Service, USA) approved pharmaceutical maize field release currently requires a separation distance of 1 mile (5280 feet or 1609 km) from the nearest maize plant (BRS, 2006). This physical isolation is eight times greater than the distance required for production of foundation maize seed or identity preservation. Second, confinement strategies for transgenic maize produced under permit include conditions that prevent transgenic maize pollen from pollinating surrounding fields. Current APHIS rules stipulate that non-pharmaceutical maize may be grown within one half mile (2640 feet) of the test site if the regulated pharma-maize is control-pollinated by detasseling or bagging procedures. For temporal isolation, pharma-maize must be planted no less than 28 days before or after the surrounding non-pharma maize is planted (APHIS, 2008).

In addition to spatial and temporal isolation, it is also possible to prevent contamination from transgenic pollen drift by producing vaccines in maize varieties that are male-sterile. A number of cytoplasmic male-sterile varieties producing near 0% viable pollen can be used for this purpose (Levings III, 1993; Gabay-Laughnan and Laughnan, 1994). Because transgenic male-sterile maize does not produce viable pollen, its female flowers can be pollinated by non-transgenic maize pollen to set seed. According to Ripplinger et al (2009), if pollen produced from pharmaceutical maize is controlled, the risk of accidental release and food supply contamination is very close to the target tolerance level of zero.

Here we describe two strategies that can be used to produce male-sterile transgenic maize lines suitable for open-field pharmaceutical production. One strategy is to directly introduce the transgenes into transformable male-sterile germplasm. A fertile and transformation-amenable Hi II genotype (Armstrong et al., 1991) was converted into a male-

sterile, transformation-amenable group of lines by crossing to cytoplasmic male sterile (CMS)-T (Texas) maize. This is the first report of successful transformation of such male-sterile lines using the biolistic-(gene gun) or *Agrobacterium*-mediated methods.

The second strategy involves transfer of the subunit vaccine gene from a fertile transgenic line to a *cms-T* B37 background. After six seasons of breeding, male-sterile transgenic maize seed can be used for open-field production using a non-transgenic pollen donor to produce seeds for large scale recombinant protein recovery, 100% of which contain the subunit vaccine gene.

Materials and Methods

Germplasm

Development of the tissue culture amenable (T) germplasm used in this study was previously described (Wise et al., 1999b). Briefly, T-cytoplasm A188 [(T) A188] was pollinated by N-cytoplasm Hi II (high Type II, or HT), the standard germplasm for maize transformation (Armstrong and Green, 1985; Armstrong et al. 1991). (N) HT is originally a tissue-culture line derived from A188 and B73, hence the selection of (T) A188 as the T-cytoplasm female. Immature embryos were excised from the resulting F_1 seed, cultured on N6E30 medium (Frame et al., 2000), and selected based on Type II callus response. Callus was then transferred to regeneration medium (Frame et al., 2000), and T cytoplasm plantlets were grown to maturity. Since the resulting ears were T cytoplasm, they were pollinated by (N) HT. Male-sterile, (T) HT plants were maintained by pollinating with the (N) HT stock (Pei, 2000). Genotype nomenclature and crosses made for seed increase and embryo production for transformation purposes are detailed in Table 1. For introgression

experiments, *cms-T* B37 seed was increased via isolation plot in our 2000 Ames, IA summer nursery.

Plant materials

For genetic transformation: Immature embryos (1.5-2.0 mm) were aseptically dissected from 10-13 day old field or green house-grown ears of male-sterile (T) and male-fertile (N) germplasm. Male-sterile phenotypes were assessed for each donor plant used or transgenic R_0 and R_1 plant produced in this study by monitoring the presence or absence of pollen shed from mature plants. Structural anatomy of (T) and (N) cytoplasm transgenic plants (Figure 1) was compared and imaged using an Olympus SZ stereo-microscope (Leeds Precision Instruments, Minneapolis).

DNA plasmid constructs and Agrobacterium strain for plant transformation

Construct pAHC25 (Christensen and Quail, 1996), containing the maize ubiquitin promoter driving both *bar* and *gus*-intron gene cassettes (Pubi-*bar*/Pubi-*gus* intron), was used for all biolistic gun-mediated transformation experiments.

Agrobacterium tumefaciens strain EHA101 (Hood et al., 1986) harboring the standard binary vector pTF102 (Frame et al., 2002), was used in all *Agrobacterium*-mediated transformation experiments. This vector carries a P35S-*bar*/P35S *gus*-intron cassette. A fresh bacteria culture was initiated from -80°C every four weeks on solid YEP (An et al., 1988) containing antibiotics (100 mg/L spectinomycin, 50 mg/L kanamycin, 25 mg/L chloramphenicol) and stored at 4°C. Cultures for weekly experiments were initiated from this 4°C "mother" plate and grown at 19°C for 3 days before use. Final OD_{550} in all experiments ranged from 0.3-0.4. All other bacteria manipulations, maintenance and pre-infection preparation steps were identical to the protocol detailed in Frame et al. (2002).

Culture media

Biolistic transformation: All media components and preparation steps were detailed in Frame et al. (2000) as adapted from Songstad et al. (1996). Petri plates (100x15 mm) were used for all bombardment and selection media.

Agrobacterium-mediated transformation: All N6 media (Chu et al 1975) components were described in Frame et al. (2002). Co-cultivation media was modified from Zhao et al. (2001) to contain 300 mg/L cysteine; resting and selection media contained a combination of cefotaxime (100 mg/L) and vancomycin (100 mg/L) for counter selection. All solid media tissue culture steps used 100x25 mm Petri-plates and all stocks were prepared as described in Frame et al. (2002).

Regeneration Media: Independent, bialaphos resistant Type II callus events generated from either transformation protocol were sub-cultured to a Petri-plate (100x25 mm) containing MS salts (Murashige and Skoog, 1962) and modified MS vitamins (Frame et al 2006), 6% sucrose, 100 mg/L myo-inositol, no hormones (Armstrong and Green, 1985), 0.3% gelrite, pH 5.8 and supplemented with 4 mg/L glufosinate (Sigma, St. Louis). Cefotaxime (250 mg/L), added to this media after autoclaving, was included for *Agrobacterium*-derived putative events. Three weeks later, mature somatic embryos were germinated in the light on MS Salts and modified MS vitamins, 3% sucrose, 100 mg/L myo-inositol and 0.3% gelrite (pH 5.8) without glufosinate or cefotaxime. Petri-plates were wrapped with vent tape.

Transformation protocols

Biolistic-gun mediated: Transformation and selection procedures were as previously described for immature zygotic embryo transformation (Frame et al., 2000) except that DNA quantity per bombardment was reduced and embryos were bombarded 3 rather than 4 days

after dissection in the present study. This protocol is used for routine production of transgenic (N) HT maize in our laboratory. Briefly, 30 zygotic embryos per plate were dissected, scutellum side up, in a 2 cm^2 grid on the middle of a Whatman No 5 filter paper laying on the surface of callus initiation medium. These plates were incubated at 28°C (dark) for 3 days. The morning of bombardment, an osmotic pretreatment of targeted embryos was begun by carefully transferring each grid of embryos (still on the filter paper) to the center of a Petri plate (100x15 mm) containing maintenance media supplemented with 0.2 M sorbitol and 0.2 M mannitol (Vain et al 1993). Four hours later, each plate of embryos on osmotic medium was bombarded once at 650, 6 cm target, ¼ inch gap. For bombardment, 10 macro-projectiles (10 µl each) were prepared from one tube of 0.6 micron gold particles (3 mg) coated with 0.075 µg plasmid pAHC25. The following morning, embryos were transferred, scutellum side up, to the surface of maintenance medium (28°C, dark) and 10 days later sub-cultured to selection medium (28°C, dark) containing 2 mg/L bialaphos. Embryos were subcultured to this medium a second time three weeks later. Six to 10 weeks after bombardment, each bialaphos resistant putative transgenic callus event was harvested to its own plate of selection medium for bulking up and regeneration as described above.

Agrobacterium-mediated: Infection, co-cultivation and selection procedures were all according to Frame et al. (2002). This protocol is used routinely in our laboratory for production of transgenic maize using an *Agrobacterium* standard binary vector to transform (N) HTF_2 germplasm. Briefly, embryos from a single ear were dissected to an eppendorf tube (2 ml) filled with bacteria-free infection medium and washed once with the same before infecting with diluted, pre-cultured (2-5 hours) *Agrobacterium* suspension. After 5 minutes infection (without vortexing), embryos were plated scutellum side up on 4-day old co-

cultivation media containing 300 mg/L cysteine, and cultured at 20°C (dark). After 3 days co-cultivation, all embryos were transferred to resting media containing cefotaxime (100 mg/L) and vancomycin (100 mg/L) but no bialaphos and incubated at 28°C (dark) for 7 days after which all embryos were transferred to selection medium containing 1.5 mg/L bialaphos and antibiotics. Selection was enhanced to 3 mg/L bialaphos for 2 two-week intervals after which individual, transgenic events could be seen emerging as actively growing Type II callus lobes on otherwise brown and dying zygotic embryo explants (Figure 2c). At this time, each putative event was sub-cultured to its own plate of selection medium. Putative events that continued to produce vigorous, embryogenic callus on selection medium containing 3 mg/L bialaphos were regenerated for 3 weeks (25°C, dark) on media containing 4 mg/L glufosinate, germinated in the light, and grown to maturity in the green house as previously described (Frame et al., 2002).

Histochemical GUS assays of callus and leaves from R_1 progeny

GUS histochemical assays (Jefferson et al, 1987) were carried out on all bialaphos-resistant putative transgenic calli recovered from selection, and on leaf pieces of male-sterile and control R_1 progeny plants to confirm expression of the *gus* transgene in segregating offspring of these germplasms. A single, 2 cm leaf piece was cut from each progeny plant, grouped by event in 60x20 mm Petri plates, and submerged in 4 ml X-Gluc solution. Plates were vacuum-infiltrated for 10 minutes at 15 in Hg then wrapped with parafilm and incubated at 28°C for 20 hours. Chlorophyll de-staining was done using a 6 hour 70% ethanol wash followed by an overnight 95% ethanol wash. A stereo-microscope was used to identify the number of blue and non-blue staining leaves for each segregating event.

R1 progeny screening for bar gene expression

After sampling progeny for GUS assays, R_1 progeny of male-sterile and control germplasm were screened for *bar* gene expression by spraying 9 day old seedlings with a 250 mg/L solution of glufosinate prepared from the herbicide Liberty® (Bayer, Research Triangle Park, NC). The herbicide resistant/sensitive scores were recorded 4 days after the spray.

Results and Discussion

Biolistic gun transformation of male-sterile germplasm

Field and green house derived immature embryos of (N) HTF_2 and (T) male-sterile germplasms were targeted for biolistic transformation in 2003 and 2004 (field only). Non-transgenic (T) F_2, and (T) $B(BC_n)$ embryo donor ear phenotypes were similar, with up to 200 zygotic embryos per ear (Figure 2a), however (T) $A(BC_n)$ embryo donor ears were consistently small with poor seed set (not shown). Transient GUS expression in (T) germplasm embryos was observed (Figure 2b), and the average transformation frequency (TF) for the (T) germplasms was equal to or higher than for the (N) HTF_2 control material (Table 2). Across the 2003 and 2004 field seasons, the highest average TF was achieved for (T) $B(BC_0)$ material (19%) compared with the (N) HTF_2 control (12%). Histochemical GUS expression was observed in 62% of the bialaphos resistant callus events assayed in this study and expression of the unselected gene did not appear to depend on germplasm (Table 2).

Except for the male-sterile tassel phenotype, all (T) germplasm transgenic plants resembled those of (N) germplasm control plants and transgenic ear sizes were similar (Figure 2e). Average female fertility (events with > 50 kernels) for transgenic events derived

from (T) germplasm was higher (71%) than for (N) HTF_2 control events (52%, Table 2). Finally, none of the 97 (T) cytoplasm transgenic plants (representing 67 independent transgenic events) taken to maturity shed pollen, while an average of 78% of the 35 (N) HTF_2 transgenic events in this study shed pollen at maturity (Table 2, Figure 1). Furthermore, all 32 R_1 progeny plants (representing 16 gun-derived events) taken to maturity after progeny tests were male sterile (not shown).

Agrobacterium-mediated transformation of male-sterile germplasm

Greenhouse and field immature embryos of (N) HTF_2 and (T) germplasms were transformed with the *Agrobacterium* standard binary vector pTF102 (EHA101) (Frame et al., 2002) in the spring and summer, respectively, of 2004. On average, the three (T) male-sterile lines transformed at higher frequencies than did the (N) HTF_2 male-fertile control, demonstrating that this T-cytoplasm material is readily transformable using the *Agrobacterium*-mediated method described (Table 3). In field 2004 experiments, a TF of 22% was achieved using (T) A(BC_1) germplasm compared with 7% for the (N) HTF_2 control. (N) HTF_2, (T) F_2 and BC_0 of (T) A and (T) B were all targeted in both the greenhouse and field experiments in 2004 and the average TF for field-derived embryos (6%) was higher than that for the greenhouse study (2%). In histochemical GUS assays carried out on all bialaphos resistant callus events recovered from *Agrobacterium*-mediated experiments, 90% of events also expressed the *gus* gene (Table 3, Figure 2d). An average of 72% of the 56 (T) male-sterile transgenic plants (representing 41 independent transgenic events) grown to maturity in the green house were female-fertile and none shed pollen. In contrast, the five (N) HTF_2 control events (1 plant/event) in this study shed pollen at maturity and all were female-fertile (Table 3).

Progeny analysis in biolistic of transgenic events

Analysis for expression of the *bar* and *gus* trangenes in gun-derived R_1 progeny plants demonstrated Mendelian inheritance of the *bar* gene in 14 of 15 (T) germplasm derived events and 2 of 3 (N) HTF_2 events (Table 4). In contrast, *gus* marker gene expression in segregating progeny did not follow the predicted pattern for Mendelian inheritance at a single locus in almost half of the events (8 of the 18 events). These abnormal segregation ratios were observed in both male-sterile and male-fertile genetic backgrounds (Table 4). Expected segregation ratios were observed for *bar* gene expression in all R_1 progeny plants derived from *Agrobacterium*-mediated transformation events (Table 4), and only 2 of these 12 events segregated abnormally for the *gus* gene.

Introgressing of transgenic maize expressing subunit vaccine into male-sterile germplasm via breeding

One strategy proposed for the safe production of pharmaceutical maize seed is to express the pharmaceutical proteins in a male sterile genetic background (Lamkey, 2004). A breeding scheme using 3 genotypes was designed for transferring the transgene into a T-cytoplasm male-sterile (CMS) background. In order for a T-cytoplasm line to produce pollen, it must carry dominant alleles of two nuclear genes, Rf_1 and Rf_2 (Schnable and Wise, 1998; Wise et al., 1999). Plants with the genotype Rf_1/rf_1, Rf_2/rf_2 can also be fertile as they carry the required copy of the dominant Rf_1 allele plus 1 copy of the dominant Rf_2 allele.

The first line, designated (T) B37 rf_1/rf_1, Rf_2/Rf_2, *b/b*, was inbred B37 that contained male-sterile cytoplasm, was male-sterile, and did not contain the LT-B gene (designated by *b/b*). The second line, designated (N) B37 rf_1/rf_1, Rf_2/Rf_2, *b/b,* contained normal cytoplasm, was male-fertile, and did not contain the LT-B gene. This line is referred to as a maintainer

line because it is needed to produce seed from the male-sterile line. The third line, designated (N) HT *rf_1/rf_1, Rf_2/Rf_2, B/B*, carried the transgene, and was male-fertile. In this line, the transgene cassette contains a gene encoding the B-subunit of the *E. coli* heat labile enterotoxin (LT-B) under control of the 27 kD gamma zein promoter (an endosperm-specific promoter). The transgene was introduced into (N) HT germplasm (Chikwamba et al., 2002a) and the LT-B expressing homozygous line was generated after several generations of self pollination. B is the transgene we wanted to transfer to T cytoplasm, and throughout this study, we assumed that HT germplasm was a non-restoring genotype

Figure 3 illustrates the breeding strategy used for transferring the LT-B gene into male-sterile (T) B37. Note that females are always listed first in a cross. In Steps 1 and 2, we generated a BC_1 population that contained 50% heterozygous transgenic seeds (N) $B37^2$ *rf_1/rf_1, Rf_2/Rf_2, B/b* and 50% non-transgenic segregants (N) $B37^2$ *rf_1/rf_1, Rf_2/Rf_2, b/b*, by crossing the male-fertile restorer line (N) B37 *rf_1/rf_1, Rf_2/Rf_2, b/b* and transgenic heterozygous line (N) HT *rf_1/rf_1, Rf_2/Rf_2, B/b* (produced in Step 1).

In Step 3, any non-transgenic segregants were effectively eliminated at the seedling stage by spraying with herbicide Liberty®. Pollen from herbicide resistant heterozygous transgenic seeds (N) $B37^2$ *rf_1/rf_1, Rf_2/Rf_2, B/b* was used to cross with the male-sterile (T) B37 *rf_1/rf_1, Rf_2/Rf_2, b/b*. This step resulted in a male-sterile population containing 50% heterozygous transgenic seeds, (T) $B37^3$ *rf_1/rf_1, Rf_2/Rf_2, B/b*, and 50% non-transgenic segregants, (T) $B37^3$ *rf_1/rf_1, Rf_2/Rf_2, b/b*.

Step 4 (a, b and c) demonstrates how we selected and bred a homozygous transgenic maintainer line (N) $B37^2$ *rf_1/rf_1, Rf_2/Rf_2, B/B* from the BC_1 maintainer line harvested in Step 2. Once the homozygous transgenic maintainer line was achieved, it was used in Step 5 to

pollinate the heterozygous sterile BC_2 generated from Step 3. Seeds from Step 5 were male-sterile and all carried the LT-B transgene, 50% (T) $B37^4$ *rf_1/rf_1, Rf_2/Rf_2, B/B* and 50% (T) $B37^4$ *rf_1/rf_1, Rf_2/Rf_2, B/b*. This seed was then used in field production (Production 1) for LT-B maize, in which a non-transgenic line was used as the pollen donor. After this four-season breeding cycle, 75% of the seeds harvested were expected to express the LT-B protein. To achieve 100% LT-B seed from field production (Production 2), the homozygous transgenic sterile line (T) $B37^5$ *rf_1/rf_1, Rf_2/Rf_2, B/B*) was generated over an additional two growing seasons (Step 6).

This breeding strategy was accomplished using both greenhouse facilities and field nurseries. Open-field containment procedures followed strict USDA-APHIS regulations. Transgenic pollen remained contained at all times by bagging tassels, or by using male sterile material. Tassel production and pollen viability in field conditions were monitored after the transgene was transferred into the male-sterile line. All of these plants produced tassels, 75% of which did not shed pollen. In the 25% of transgenic male-sterile plants that did shed pollen, no viable pollen was detected (X. Xu, unpublished), indicating that no fertility was restored in these plants.

Conclusions

We describe possible strategies for using cytoplasmic male sterile (*cms*-T) germplasm in open-field production of transgenic maize producing a pharmaceutical product. In the first approach, a male-sterile germplasm amenable for tissue culture and transformation was developed from the widely used Hi Type II genotype. Male-sterile transgenic events (non-pharmaceutical) were generated using the biolistic gun and *Agrobacterium*-mediated

transformation methods and grown to maturity in the greenhouse. While *in vitro* culture induced variation did not alter the male sterile phenotype in this study, the possibility that this could occur cannot be ruled out.

Secondly, a pharmaceutical transgene introduced into (N) cytoplasm using the biolistic gun was transferred into a male-sterile genotype via conventional breeding. Our transgene or the transgene insertion site did not restore male fertility in transgenic maize plants (although a non viable pollen was shed in 25% of plants), thereby affirming the applicability of this technology. However, it is prudent to state that any potential restoration or reversion should be examined on a gene-by-gene basis.

There may be a third strategy that could accelerate production of male-sterile homozygous pharmaceutical maize seed. Both male-fertile and male-sterile genotypes can be directly transformed with the transgene of interest. The male-fertile transgenic maize that serves as the maintainer line could be made homozygous and its pollen could then be used to cross to the male-sterile transgenic female. Progeny would then carry two copies of the transgene at different loci. While this practice may save 2-3 seasons of breeding, the presence of multiple transgene copies may cause transgene silencing (Meng et al., 2003; Shou et al., 2004).

The outcome of using *cms*-T maize for large scale open-field release of a pharmaceutical product is untested. T-cytoplasm maize is highly sensitive to the host selective pathotoxin HmT, produced by *Cochlobolus heterostrophus* Drechsler race T (Wise et al. 1999a). While no obvious pathogen damage was observed in the small field experiments carried out in our study, it is important to bear in mind that recurrent large scale

field release of *cms*-T genotype may encourage reappearance of Southern maize leaf blight, which could lead to yield loss in Texas male-sterile, pharmaceutical maize.

Acknowledgements

The authors thank Deqing Pei for initiating the Hi II T-cytoplasm lines for transformation, Tina Paque for greenhouse and field assistance in preparing transformation starting materials, Lise Marcell, Marcy Main, Jenifer McMurray and Kyle Taylor for their assistance in transformation experiments and progeny tests, Horan BioProduction, Kent Berns, Kevin VanDee, Mark Honeyman for the breeding program. Partial financial support for this project is from Iowa State University Plant Science Institute BioPharmaceuticals and BioIndustrials Initiative, USDA-ARS CRIS project 3625-21000-035-00D, and Pioneer Hi-Bred International.

References

An G, Ebert PR, Mitra A, Ha SB (1988) Binary vectors. *In* SB Gelvin, RA Schilperoort, eds. Plant Molecular Biology Manual. Kluwer Academic Publishers, Dordrecht, The Netherlands

APHIS (2008) Guidance for APHIS Permits for Field Testing or Movement of Organisms Intended for Pharmaceutical or Industrial Use. http://www.aphis.usda.gov/brs/pdf/Pharma_Guidance.pdf.

Armstrong CL, Green CE, Phillips RL (1991) Development and availability of germplasm with high Type II culture formation response. Maize Genet Coop Newsletter **65**: 92-93

Armstrong CL, Green CE (1985) Establishment and maintenance of friable, embryogenic maize callus and the involvement of L-proline. Planta **164**: 207-214

Bailey MR, Woodard SL, Callaway E, Beifuss K, Magallanes-Lundback M, Lane JR, Horn ME, Mallubhotla H, Delaney DD, Ward M, Van Gastel F, Howard JA, Hood EE (2004) Improved recovery of active recombinant laccase from maize seed. Appl Microbiol Biotechnol **63:** 390-397

BRS (2006) Permitting Genetically Engineered Plants That Produce Pharmaceutical Compounds. In: Biotechnology Regulatory Service Fact Sheet. February 2006. http://www.aphis.usda.gov/publications/biotechnology/content/printable_version/BRS_FS_pharmaceutical_02-06.pdf

Chikwamba R, Cunnick J, Hathaway D, McMurray J, Mason H, Wang K (2002) A functional antigen in a practical crop: LT-B producing maize protects mice against

Escherichia coli heat labile enterotoxin (LT) and cholera toxin (CT). Transgenic Res **11:** 479-493.

Chikwamba R, McMurray J, Shou H, Frame B, Pegg SE, Scott P, Mason H, Wang K (2002) Expression of a synthetic E. coli heat-labile enterotoxin B sub-unit (LT-B) in maize. Mol. Breeding **10:** 253-265

Chu CC, Wang CC, Sun CS, Hsu C, Yin KC, Chu CY, Bi FY (1975) Establishment of an efficient medium for anther culture of rice through comparative experiments on the nitrogen source. Sci. Sinica **18**: 659-668

Christensen AH and Quail PH (1996) Ubiquitin promoter-based vectors for high-level expression of slectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res **5**: 213-218.

Frame B, Zhang H, Cocciolone S, Sidorenko L, Dietrich C, Pegg S, Zhen S, Schnable P, Wang K (2000) Production of transgenic maize from bombarded Type II callus: effect of gold particle size and callus morphology on transformation efficiency. In Vitro Cell Dev Biol- Plant **36**: 21-29

Frame BR, Shou H, Chikwamba RK, Zhang Z, Xiang C, Fonger TM, Pegg SEK, Li B, Nettleton D, Pei D, Wang K (2002) *Agrobacterium-tumefaciens*-mediated transformation of maize embryos using a standard binary vector system. Plant Physiology 129: 13-22.

Frame BR, McMurray JM, Fonger TM, Main ML, Taylor KW, Torney FJ, Paz MM, Wang K (2006) Improved *Agrobacterium*-mediated transformation of three maize inbred lines using MS salts. Plant Cell Reports, **25**: 1024-1034.

Gabay-Laughnan S, Laughnan JR (1994) Male sterility and restorer genes in maize. *In* M Freeling, V Walbot, eds, The Maize Handbook. Springer, New York, NY, USA, pp 418-423.

Hood EE, Helmer GL, Fraley RT, Chilton MD (1986) The hypervirulence of *Agrobacterium tumefaciens* A281 is encoded in a region of pTiBo542 outside of T-DNA. J Bacteriol **168**: 1291-1301.

Hood EE, Kusnadi A, Nikolov Z, Howard JA (1999) Molecular farming of industrial proteins from transgenic maize. Adv Exp Med Biol **464:** 127-147

Hood EE, Witcher DR, Maddock S, Meyer T, Baszczynski C, Bailey M, Flynn P, Register J, Marshall L, Bond D, Kulisek E, Kusnadi A, Evangelista R, Nikolov Z, Wooge C, Mehigh RJ, Hernan R, Kappel WK, Ritland D, Li CP, Howard JA (1997) Commerical production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. Molecular Breeding **3:** 291-306

Jefferson RA (1987) Assaying chimeric genes in plants. The *gus* gene fusion system. Plant Molec Biol Rep **5**: 287-405

Lamkey KR (2004) Seed production in corn and soybeans. p. 54-73. *In:* Andow, D.A. (ed.) A growing concern: Protecting the food supply in an era of pharmaceutical and industrial crops. Union of Concerned Scientists. Cambridge, MA.

Levings III, CS (1993) Thoughts on Cytoplasmic Male Sterility in cms-T Maize. The Plant Cell **5**: 1285-1290

Meng L, Bregitzer P, Zhang S, Lemaux P (2003) Methylation of the exon/intron region of the *Ubi*1 promoter complex correlates with transgene silencing in barley. Plant Mol Biol **53**: 327-340

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant **15**: 473-497

Pei, D. (2000) Masters thesis – Iowa State University: Positioning the *Rf8* and *Rf** restorers of fertility loci in the maize genome.

Ramessar K, Sabalza M, Capell T, Christou P (2008) Maize plants: an ideal production platform for effective and safe molecular pharming. Plant Science 174: 409-419

Ripplinger DG, Hayes DJ, Goggi AS, Lamkey K (2009) Insuring against losses from from transgenic contamination: the case of pharmaceutical maize. Amer. J. Agr. Econ **91:** 322-334

Schnable PS, Wise RP. (1998) The molecular basis of cytoplasmic male sterility and fertility restoration. Trends in Plant Science **3**:175-180.

Shou H, Frame BR, Whitham SA, Wang K (2004) Assessment of transgenic maize events produced by particle bombardment or *Agrobacterium*-mediated transformation. Mol Breeding **13**:201-208

Songstad DD, Armstrong CL, Petersen WL, Hairston B, Hinchee MAW (1996) Production of transgenic maize plants and progeny by bombardment of Hi II immature embryos. In Vitro Cell Dev Biol – Plant 32: 179-183

Strayer D (2002) Identy-preserved systems: A reference handbook. CRC Press. Boca Raton, FL.

Strayer D (2004) Post-harvest grain handling, storage, and shipping in corn and soybeans. p. 89-101. *In:* Andow, D.A. (ed.) A growing concern: Protecting the food supply in an era of pharmaceutical and industrial crops. Union of Concerned Scientists. Cambridge, MA.

Streatfield SJ (2007) Approaches to achieve high-level heterologous protein production in plants. Plant Biotechnol J **5:** 2-15

Streatfield SJ, Jilka JM, Hood EE, Turner DD, Bailey MR, Mayor JM, Woodard SL, Beifuss KK, Horn ME, Delaney DE, Tizard IR, Howard JA (2001) Plant-based vaccines: unique advantages. Vaccine **19:** 2742-2748.

Vain P, McMullen MD, Finer JJ (1993) Osmotic treatment enhances particle bombardment-mediated transient and stable transformation from maize cell culture. Plant Cell Rep 12:84-88

Wise RP, Bronson, CR, Schnable, PS, Horner, HT (1999a) The genetics, pathology and molecular biology of T-cytoplasm male sterility in maize. Advances in Agronomy **65**: 79-130.

Wise RP, Gobelman-Werner K, Pei D, Dill CL, Schnable P (1999b) Mitochondrial transcript processing and restoration of male fertility in T-cytoplasm maize. Journal of Heredity 90 (3): 380-385

Witcher DR, Hood EE, Peterson D, Bailey M, Bond D, Kusnadi A, Evangelista R, Nikolov Z, Wooge C, Mehigh R, Kappel W, Register J, Howard JA (1998) Commercial production of beta-glucuronidase (GUS): A method system for the production of proteins in plants. Molecular Breeding **4:** 301-312

Yang SH, Moran DL, Jia HW, Bicar EH, Lee M, Scott MP (2002) Expression of a synthetic porcine alpha-lactalbumin gene in the kernels of transgenic maize. Transgenic Res **11:** 11-20.

Zhao ZY, Gu W, Cai T, Tagliani L, Hondred D, Bond D, Shroeder S, Rudert M, Pierce D (2001) High-throughput genetic transformation mediated by *Agrobacterium tumefaciens* in maize. Mol Breeding **8**: 323-333

Zhong GY, Peterson D, Delaney DE, Bailey M, Witcher DR, Register JC, III, Bond D, Li CP, Marshall L, Kulisek E, Ritland D, Meyer T, Hood EE, Howard JA (1999) Commercial production of aprotinin in transgenic maize seeds. Molecular Breeding **5:** 345-356.

Table 1. Description of germplasm used for genetic transformation *

Base germplasm ID	Pedigree	Phenotype		Cross for seed increase	Cross for transformation	Transformation genotype nomenclature **
		Female	Male			
(T) pA	(T) A188/(N) HT pA x (N) HT pA BC_6	fertile	sterile	(T) pA x (N) HTpA	none	NT
(T) F_1	(T) A188/(N) HT pA x (N) HT pB BC_0	fertile	sterile	(T) pA x (N) HTpB	(T) F_1 x (N) HTF_1	(T) F_2
(T) PeiA(BC_0)	[F98 (T) A188/HT A/B x (N) A188/HT A/B] x (N) HT pA Tissue culture regenerated plants from the greenhouse	fertile	sterile	(T) PeiA x (N) HTpA	(T) PeiA(BC_n) x (N) HTF_1	(T) A(BC_0) in 2003/4 (T) A(BC_1) in 2004
(T) PeiB(BC_0)	[F98 (T) A188/HT A/B x (N) A188/HT A/B] x (N) HT pB Tissue culture regenerated plants from the greenhouse	fertile	sterile	(T) PeiB x (N) HTpB	(T) PeiB(BC_n) x (N) HTF_1	(T) B(BC_0) in 2003/4 (T) B(BC_1) in 2004
(N) HTpA	HT parent A	fertile	fertile	(N) HTpA x (N) HTpA	none	NT
(N) HTpB	HT parent B	fertile	fertile	(N) HTpB x (N) HTpB	none	NT
(N) HTF_1	HT pA x HT pB	fertile	fertile	(N) HTpA x (N) HTpB	(N) HTF_1 x (N)	(N) HTF_2

* Male-sterile (T) pA, (T) PeiA or (T) PeiB seed was increased in 2003 by pollinating with the male-fertile (N) HTpA or (N) HTpB or to make the F_1 hybrid cross (T) pA x (N) HTpB. Embryo production for transformation purposes was achieved by crossing these stock plants with (N) HT F_1 pollen in all cases. In 2003 and 2004, pollinations were carried out on plants grown from the original pedigree [(T) BC_0 x (N) HTF_1]. As well, in 2004 pollen from (N) HTF_1 was taken to one generation advanced (T) A(BC_1) and (T) B(BC_1) plants to produce (T) A or (T) B(BC_1) immature embryos for transformation experiments. (N) HTF_1 plants were sib or self pollinated to produce (N) HTF_2 embryos for transformation of control material in this study.

** Nomenclature of immature zygotic embryo donor plant for genetic transformation.

(T): Texas male sterile cytoplasm

(N): Normal male fertile cytoplasm

NT: not tested

Table 2. Biolistic gun transformation frenquency (TF), seed set, and pollen shed for Hi II (N) and (T) germplasms

Source	Donor plant genotype	# Ears	# Emb bb[a]	# Bialaphos resist events	% TF[b]	% Stable events GUS positive[c]	% Events with >50K[d]	% Events with pollen shed[e]
2003 gh	(N) HTF_2	15	945	93	10	65	50	87
	(T) F_2	19	1400	232	17	60	82	0
2003 field	(N) HTF_2	16	1110	136	12	61	55	70
	(T) F_2	11	600	87	15	46	58	0
	(T) A(BC_0)	13	881	116	13	54	75	0
	(T) B(BC_0)	15	930	228	25	58	70	0
2004 field	(N) HTF_2	9	800	101	13	74	NT	NT
	(T) A(BC_0)	2	272	48	18	69	NT	NT
	(T) A(BC_1)	2	138	7	5	57	NT	NT
	(T) B(BC_0)	6	382	50	13	80	NT	NT
	(T) B(BC_1)	3	81	23	28	65	NT	NT

[a] all embryos bombarded (bb) with pAHC25

[b] # independent bialaphos (2 mg/L) resistant events / total # embryos bombarded (X100)

[c] all callus events were assayed for histochemical GUS expression

[de] # independent events to seed (1 or 2 plants per event) ranged from 8 -30 events

NT (not tested)

TABLE 3 - Agrobacterium-*mediated transformation frequency, seed set, and pollen shed for Hi II (N) and (T) germplasms.*

Source	Donor plant genotype	# Ears	# Emb inf[a]	# Bialaphos resist events	% TF[b]	% Stable events GUS positive[c]	% Events with >50K[d]	% Events with pollen shed[e]
2004 gh	(N) HTF_2	2	286	5	2	100	93	100
	(T) F_2	18	1903	27	1	91	86	0
	(T) A (BC_0)	7	579	25	4	100	65	0
	(T) B (BC_0)	9	748	12	2	83	66	0
2004 field	(N) HTF_2	12	659	44	7	88	NT	NT
	(T) F_2	5	290	12	4	92	NT	NT
	(T) A (BC_0)	15	907	82	9	90	NT	NT
	(T) A (BC_1)	8	466	102	22	84	NT	NT
	(T) B (BC_0)	9	448	21	5	100	NT	NT
	(T) B (BC_1)	12	661	28	4	78	NT	NT

[a] all embryos infected (inf) with pTF102(EHA101) standard binary vector and co-cultivated on medium with 300 mg/L L-cysteine.
[b] # independent bialaphos (3 mg/L) resistant events / total # embryos infected (X100).
[c] all callus events were assayed for histochemical GUS expression.
[de] # independent events to seed (1 or 2 plants per event) ranged from 12 -23 events.
NT (not tested).

Table 4. Segregation analysis for *gus* and *bar* gene expression in R_1 generation progeny plants[a]

Genotype	Event ID	Segregation Ratio					
		Herbicide		X^{2f}	Gus		X^{2f}
		Res[b]	Sen[c]		Pos[d]	Neg[e]	
Biolistic events[g]							
HTF_2	187	7	8	0.1	6	9	1
(T) F_2	753	13	12	0.0	0	25	25
(T) F_2	771	4	21	11.6	3	22	14
(T) F_2	777	11	18	1.7	13	16	0
(T) F_2	760	12	17	0.9	0	29	29
(T) F_2	752	11	18	1.7	15	14	0
HTF_2	193	8	20	5.1	0	28	28
(T) A(BC_0)	1057	17	13	0.5	0	30	30
(T) A(BC_0)	1066	13	15	0.1	5	23	12
(T) A(BC_0)	1028	15	14	0.0	15	14	0
(T) A(BC_0)	1024	8	12	0.8	9	11	0
(T) A(BC_0)	1018	12	16	0.6	0	28	28
HTF_2	186	12	18	1.2	0	30	30
(T) B(BC_0)	1252	14	11	0.4	17	8	3
(T) B(BC_0)	1108	16	14	0.1	15	15	0
(T) B(BC_0)	1207	10	15	1.0	16	9	2
(T) B(BC_0)	1176	9	15	1.5	13	12	0
(T) B(BC_0)	1204	18	11	1.7	19	10	3
Agro events[g]							
HTF_2	63	12	15	0.3	16	11	1
(T) F_2	468	16	14	0.1	15	15	0
(T) F_2	465	13	12	0.0	12	13	0
(T) F_2	405	15	12	0.3	0	27	27
HTF_2	72	14	15	0.0	15	14	0
(T) A(BC_0)	453	16	11	0.9	14	13	0
(T) A(BC_0)	446	10	16	1.4	9	17	3
(T) A(BC_0)	426	17	12	0.9	15	14	0
HTF_2	64	20	10	3.3	15	15	0
(T) B(BC_0)	429	14	16	0.1	18	12	1
(T) B(BC_0)	444	12	15	0.3	5	22	11
(T) B(BC_0)	445	13	17	0.5	16	11	1

[a] Transgenic plants were crossed as the female parent with pollen from non-transformed B73 plants or Hi II (Agro event #'s 64, 72, 444)

[b] Res, resistant to glufosinate spray (*bar*-expresser)

[c] Sen, sensitive to glufosinate spray (*bar* non-expresser)

[d] Pos, GUS assay positive (*gus*-expresser)

[e] Neg, GUS assay negative (*gus* non-expresser)

[f] X^2=3.8 (0.05, 1 df)

[g] biolistic events transformed with pAHC25 and Agro events with pTF102 (EHA101)

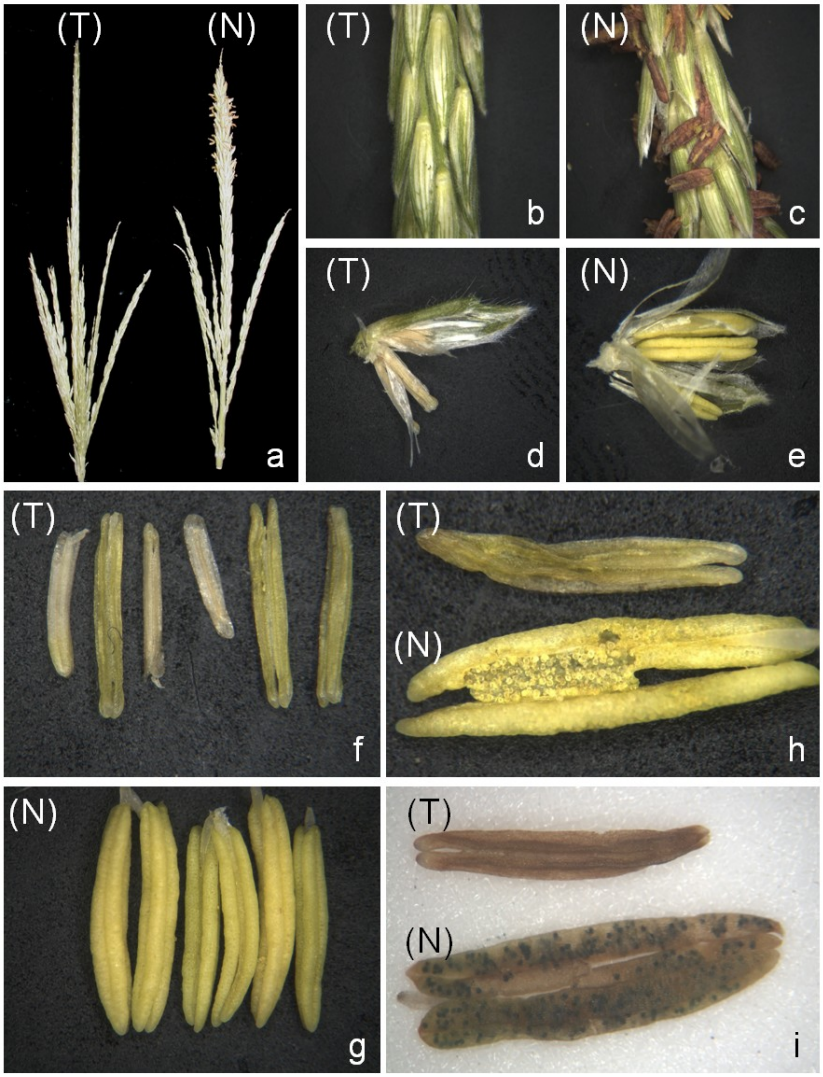

Figure 1. Tassels of gun-derived pAHC25 transgenic (T) F_2 (left) and (N) HTF_2 (right) plants at anthesis in greenhouse. No anther extrusion is visible on tassel of (T) F_2 transgenic event (b) while (N) HTF_2 event (c) is undergoing anthesis. Florets of transgenic (T) F_2 plants contain degenerated anthers (d and f) while those of (N) HTF_2 transgenic plants are filled (e) with pollen grains (h, bottom). No pollen grains are visible in (T) germplasm transgenic anthers (h, top). GUS expressing, viable pollen grains are visible in X-Gluc stained anthers of (N) transgenic germplasm (i, bottom) but not (T) transgenic germplasm (i, top) transformed with pAHC25 using the biolistic gun.

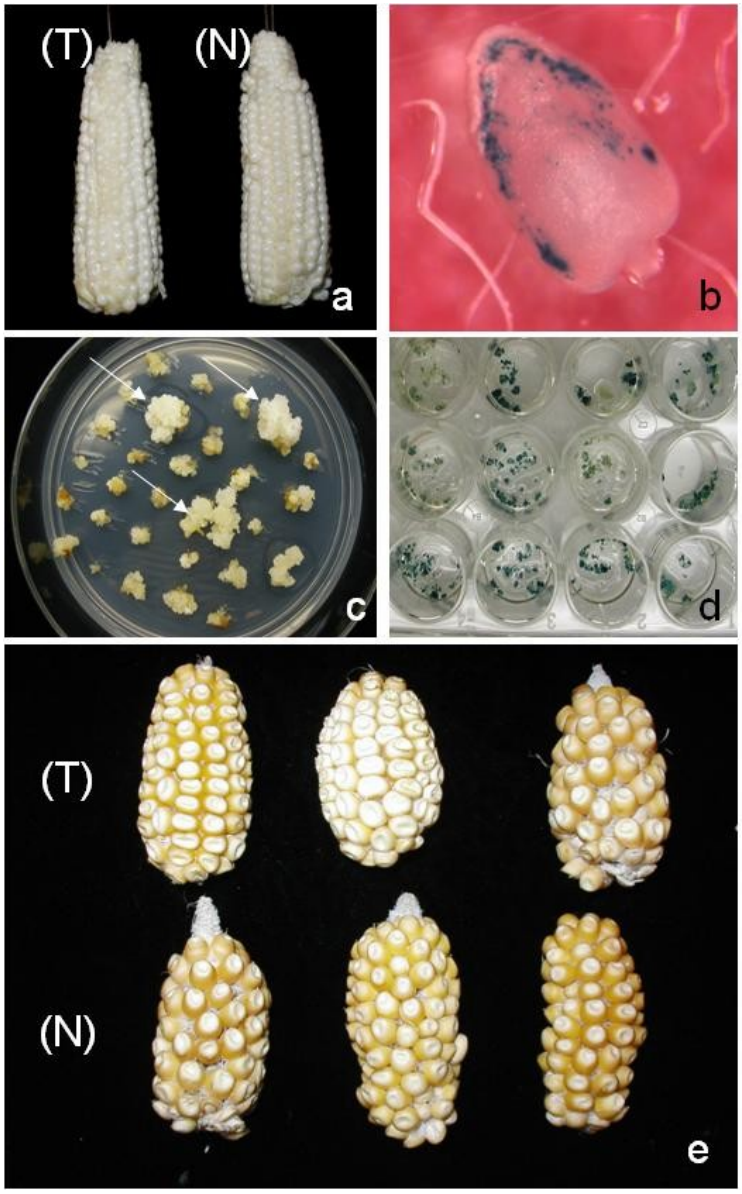

Figure 2. (N) HTF_1 (right) and (T) F_1 (left) greenhouse grown donor ears (a). (T) F_2 immature embryo stained for transient GUS expression 5 days after Agro-infection with pTF102 (b). Putative transgenic events (arrows) emerging from 3 mg/L bialaphos selection of Agro-infected (T) A(BC_1) field 2004 embryo infections (c). Stable GUS expression in putative callus events from field 2004 *Agrobacterium*-derived (N) and (T) events (d). Seed set on 3 (T) F_2 (top) and 3 (N) Hi II F_2 (bottom) pAHC25 gun-derived transgenic events (e).

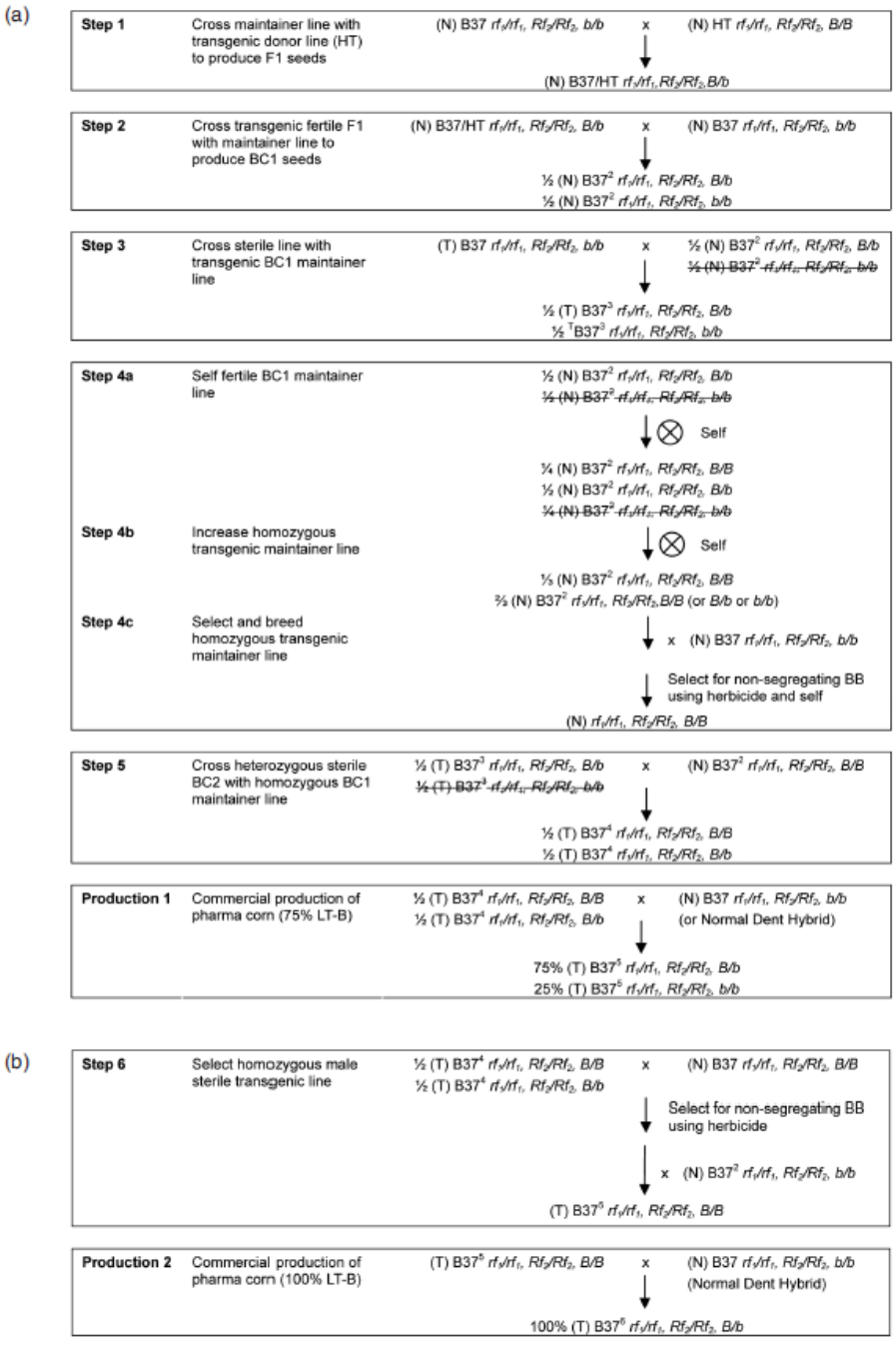

Figure 3. Breeding schemes for achieving (a) 75% and (b) 100% LT-B maize seed production in the field. T, male sterile, N, male fertile, *rf1/rf1, Rf2/Rf2*, nuclear restorer gene genotype, B, dominant transgene LT-B. Genotype with letter strikethrough represents non-transgenic segregants eliminated by herbicide spraying at plantlet stage.

CHAPTER 6: EVALUATION OF GFP EXPRESSION LEVELS IN TRANSGENIC MAIZE LINES WITH DECREASED SEED PROTEIN γ-ZEIN OR OVER-EXPRESSING A ZEIN TRANSCRIPTION FACTOR PBF

Xing Xu, Liyan Yang and Kan Wang

Abstract

Maize is an important crop worldwide and has been used to produce recombinant proteins as a safe and cost effective source. However, the low levels of recombinant protein expression challenge the application of plant-based systems. In the present study, we present two strategies for improving the recombinant protein expression in maize seeds. The first strategy is to knock down the major seed storage protein γ-zein. The 16- and 27-kDa γ-zeins were knocked down by RNA interference (γRNAi) and crossed with a 27-kDa γ-zein promoter-driving green fluorescent protein (γGFP) line. The GFP expression in F1 progeny was decreased in the γGFP/γRNAi seeds, which is opposite to what was expected. Because the γ-zeins are essential for correct protein body formation where all zeins are stored, the irregular and disrupted protein body in γRNAi seeds might have negative effects on the GFP expression driven by the 27-kDa γ-zein promoter. The second strategy was to over-express a 27-kDa γ-zein transcription factor prolamin-box binding factor (PBF) under the control of the maize ubiquitin promoter. By crossing the PBF over-expressing maize line (oxPBF) with the 27-kDa γ-zein promoter driving recombinant protein transgenic maize line, the recombinant protein expression was expected to enhance. Data collected from this work sheds light on seed protein regulation

mechanisms, which, in turn, provides information for further improvement of recombinant protein expression in seed.

Introduction

The use of plants for recombinant protein expression has many advantages over animal and microbial systems, including high yield, low cost, low contamination risk, and post-translational modification ability (Fischer and Emans, 2000; Giddings, 2001; Ma et al., 2003; Ramessar et al., 2008). Seeds have been used to produce recombinant protein, since plant seed is a natural organ for protein synthesis and storage with low protease activities (Muntz, 1998; Stoger et al., 2005). Maize is one of the major crops with high yields, not only providing nutrients to humans and livestock, but also to producing fuel ethanol. More than 90% of fuel ethanol is maize-derived (Wolt and Karaman, 2007). Thus, maize seed is an ideal system for co-production of recombinant proteins and fuel ethanol. Usually, 99% of the seed mass remains after recombinant protein recovery (Paraman et al., 2010) which can be used for ethanol production (Johnson et al., 2003). The co-production strategy not only increases the profitability of maize-derived products, but also reduces the U.S. dependence on foreign fuel and lowers green gas emissions. Many pharmaceutical and industrial proteins have been successfully produced in transgenic maize, like aprotinin (Zhong et al., 2007), the B subunit of the heat labile enterotoxin (LT-B) (Chikwamba et al., 2002), collagen (Xu et al., 2011b), gastrogenterities virus (TGEV) capsid protein (Basaran and Rodrıguez-Cerezo, 2008), and immunoglobulin G (IgG) (Lee and Forciniti, 2010).

There is a major challenge however when maize is used to produce these exogenous proteins—low expression. The first plant-derived recombinant protein was reported in 1986, which was a human growth hormone produced in transgenic tobacco (Barta et al., 1986). The first plant-derived recombinant protein for commercial purposes was reported in 1997, a maize-derived recombinant avidin—an egg protein (Hood et al., 1997). There are many factors limiting the commercialization of plant-derived recombinant protein, but low expression is a major issue both in academic and industrial areas. Numerous studies about how to enhance foreign protein production in plants have been published, including multiple strategies, such as tissue-specific expression, subcellular targeting, enhancing mRNA stability, and post-translational modification (Egelkrout et al., 2012; Hood and Vicuna Requesens, 2011; Kermode, 2006; Obembe et al., 2011; Xu et al., 2011a). Here, we evaluate two strategies for enhancing the expression of recombinant proteins in maize seed.

The first strategy is to knock out the maize seed storage proteins. Previous studies have shown plant seeds have the ability to rebalance their content after singular major proteins is silenced. In the natural maize *o2* mutant, non-zein proteins have increased, while zeins are significantly suppressed (Coleman and Larkins, 1999). High-oleic and high-stearic oil content cottonseed have been produced by the RNAi mediated knock-down of two key fatty acid desaturase genes in cottonseed (Liu et al., 2002). By suppressing β-conglycinin in soybean via sequence-mediated gene silencing (Kinney et al., 2001), the glycinin promoter-driven green fluorescent protein (GFP) expression was increased four-fold when introgressed into the β-conglycinin suppression line compared to the parent line (Schmidt and Herman, 2008).

In maize, the major seed storage proteins, zeins, make up about 70% of the total endosperm content in maize seeds (Zarkadas et al., 2000). Zeins are a mixture of alcohol-soluble proteins classified into four sub-families by their structure—α-, β-, γ-, and δ-zeins (Esen, 1987). These four subfamilies can be further classified as: 1) 19- and 22-kDa α-zeins; 2) 50-, 16-, and 27-kDa γ-zeins; 3) 15-kDa β-zein; and 4) 10- and 18-kDa δ-zeins (Sodek and Wilson, 1971; Xu and Messing, 2008). In the present study, the γ-zein protein was selected to be silenced by RNAi, since the promoter for the 27-kDa γ-zein is one of the strongest promoters among zeins (Ueda and Messing, 1991) and is one of the most popular seed specific promoters used for recombinant production in maize. A transgenic maize line expressing GFP under the control of 27-kDa γ-zein promoter (γGFP) was also generated and crossed with the γ-zein knock-down line (γRNAi). Here, we hypothesize the GFP expression may be enhanced in the γ-zein knock-down background due to maize seed protein rebalancing.

The second strategy was to enhance the expression of recombinant proteins by co-expressing of a transcription factor. In plants, 5-7% of protein encoding genes are regulated by transcription factors (Shiu et al., 2005). For example, in *Arabidopsis thaliana*, PAP1/PAP2 (R2R3-MYB) regulate anthocyanin biosynthesis (Borevitz et al., 2000) and AtMYB12/AtMYB11/AtMYB111 regulate flavonol accumulation (Mehrtens et al., 2005; Stracke et al., 2007). In maize, the prolamin-box binding factor (PBF) has a *trans*-activation on grain storage protein genes (Mena et al., 1998; VicenteCarbajosa et al., 1997). The expression of GUS (beta-glucuronidase, a reporter gene) was increased by seven-fold with co-transfection of a PBF over-expression vector 24 hours after bombardment in developing maize seeds (from 8 to 30 days after pollination (DAP))

(Marzabal et al., 2008). GUS is driven by the γ-zein promoter and PBF is driven by the 35S promoter, which indicated the PBF activated the expression of a recombinant gene under the control of the γ-zein promoter. A recent study (Wu and Messing, 2012a) also showed a similar result—the expression of the 27-kDa γ-zein promoter driving GFP was enhanced by co-expressed with the 35S promoter driving PBF. The enhanced GDP expression was observed in 10 DAP embryos 24 hours after infection or calli and root tissues after regenerated on regeneration medium. However, no report on the effects of over-expression of PBF on recombinant protein expression in stably transformed maize lines is available. In the present study, PBF was over-expressed in maize and to be crossed with a γ-zein promoter driven recombinant protein (GFP or collagen). We hypothesize the expression of recombinant proteins driven by the γ-zein promoter may be increased because of the up-regulation of PBF. The objective of this study is to evaluate whether any changes in endogenous protein contents or alteration of endogenous seed protein gene expression may enhance expression of recombinant proteins in maize seeds.

Materials and Methods

Vector construction

Two constructs were designed (Figure 1): 1) pX006 generating a double-stranded RNA to knock down the native γ-zeins in maize seeds and 2) pX012 over expressing the prolamin-box binding factor (PBF) in maize. The construct, pX006, consists of a 478 bp fragment of the 672 bp 27-kDa γ-zein full coding sequence (NCBI GI#: 16305108) amplified with Gateway primers Gat8-07 (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTACC-GGTTCATCTACCGCCG-3')

and Gat8-08 (5'-GGGGACAAGTTTGTACAAAAAAGCA-GGCTTACCTGGCCGCTTTGCGGCT-3') by polymerase chain reaction (PCR). The PCR was performed using the following conditions for 25 cycles—30 s denaturation at 95°C, 30 s annealing at 72°C, and 3 min extension at 68°C. A typical PCR reaction consists of 100 ng of template DNA, 0.8 mM of dNTPs, 2 mM of $MgCl_2$, Taq DNA polymerase buffer, and 0.5 U Taq DNA polymerase (Bioline USA Inc, Taunton, MA) in a final volume of 25 μL. The amplified 27-kDa γ-zein fragment was cloned into the pB7GWIWG2II0 (Karimi et al., 2002) backbone vector via DNA recombination using MultiSite Gateway® Three-Fragment Vector Construction Kit (cat # 12537-023, Life Technologies, USA). The 27-kDa γ-zein fragment was inserted into two positions on the vector with one in a reverse direction and the other in a forward direction, separated by a 1355 bp intron containing a 703 bp chloramphenicol resistance gene (Cm^R) (Figure 1, pX006) (Close and Rodriguez, 1982; Karimi et al., 2002). The reverse-intron-forward DNA structure driven by a promoter of a maize seed storage protein 27-kDa γ-zein should form a double-stranded RNA after transcribed, and trigger the RNA interference (RNAi) process for both 27 and 16-kDa γ-zeins in maize, since these two genes are highly conserved in DNA sequence. The pX012 construct consists of the full length of the PBF coding sequence (NCBI GI#: 162458993), including the 5' and 3' UTR (untranslated region) driven by a constitutive promoter from the maize ubiquitin gene (Figure 1, pX012). Both pX006 and pX012 are binary vectors in which transgenes can be delivered into plant cells with a bar gene cassette as a selection marker via the *Agrobacterium*-mediated transformation method.

Production of transgenic plants

The maize transformation work was completed by the Plant Transformation Facility at Iowa State University as described (Frame et al., 2002). In short, constructs, pX006 and pX012, were introduced into immature embryos of HiII maize genotype by using an *Agrobacterium*-based transformation method. Herbicide-resistant calli were selected and analyzed by PCR for the presence of the transgene. PCR-positive calli were regenerated into seedlings. The plants were brought to maturity in the greenhouse at the Plant Transformation Facility. The transgenic lines were named as A222 and A339 for constructs pX006 and pX012, respectively.

Two events selected from a total of fourteen events of A222 (γRNAi) line were crossed with one transgenic maize line (P309) carrying green fluorescent protein (GFP) under the 27-kDa γ-zein promoter (γGFP) previously generated (Moeller et al., 2009). The T_2 seeds of γGFP/γRNAi were harvested, analyzed, and advanced to T_3 by self-pollination. The A222 (γRNAi) line was also crossed with the A339 (oxPBF) line. Four genotypes of the seed were generated in the T_2 progeny—+/+; +/+, oxPBF/+; γRNAi/+, oxPBF/+; +/+, and +/+; +/γRNAi.

PCR analysis of transgenic plants

Total genomic DNA was extracted by using the cetyl trimethyl ammonium bromide (CTAB) method (Murray and Thompson, 1980) from maize calli or seeds. The presence of the transgene was verified by PCR. To detect pX006, PCR was performed under the following conditions for 35 cycles by primers x8-15 (5'-GCAAGATGTGGCGTGTTACG -3') and x8-16 (5'- GGTGGCACTTG-

TTGGTATGAGACT -3') designed for amplifying a fragment of the intron: 30 s denaturation at 95°C, 15 s annealing at 55°C, 30 s, and extension at 70°C. To detect pX012, PCR was performed under the following conditions by primers x11-03 (5'-GAGTTTAAGATGGATGGAA-3'), within the region of the ubiquitin promoter, and x11-04 (5'- GTAGCGTGGTT-GTGACA -3'), within the region of the PBF coding sequence, for 30 cycles: 30 s denaturation at 95°C, 30 s annealing at 55°C, and 20 s extension at 70°C. PCR products were analyzed on 1% agarose gel. The expected sizes of PCR products are 332 bp (pX006) and 610 bp (pX012). The gel was stained with ethidium bromide (0.5 μg/ml) for 20 min. The product sizes were determined by using 1 kb DNA Ladder (cat # N3232S, New England Biolabs).

Seeds imaging

The opaque phenotype was observed in the transgenic line A222. The mature seeds from the A222 line and wide type B73 were imaged with a digital camera on the top of a white light transilluminator (cat # TW-43, UVP LLC.).

Protein extraction and Coomassie Brilliant Blue staining

Total zein proteins (alcohol-soluble) from maize seed were extracted by using ethanol-based extraction buffer consisting of 70% ethanol, 61 mM sodium acetate, and 100 nm DL-Dithiothreitol (DTT). Seeds were ground and 100 mg powder was added to 1.0 mL of extraction buffer in a 1.5 mL Eppendorf tube. The tube was laid horizontally on a shaker at 225 rpm at 37^{o}C for 1 hour. Then, the tube was spun 10 minutes in a micro-centrifuge.

Two micro liters of the total zein proteins extracted from the A222 seed powder were lyophilized overnight and re-dissolved in 30 µL Laemmli sample buffer (cat # 161-0737, Bio-Rad). The sample was loaded onto a 12 % polyacrylamide SDS-PAGE gel (cat # 465-1043, Bio-Rad). Electrophoresis was performed at 200 V for 35 min. The gel was stained with a Bio-Safe Coomassie Stain (cat # 161-0786, Bio-Rad) for 20 min and de-stained overnight in deionized water.

The total water-soluble protein (TSP) from maize seeds was extracted by using the following protein extraction buffer at a ratio of 10 µL buffer per milligram of seed powder for 2 hours at 37 ^{o}C: 25 mM sodium phosphate (pH 6.6), 100 mM NaCl, 0.1% Triton X-100 (v/v), 1 mM EDTA, 10 µg/ml of leupeptin, and 0.1 mM serine protease inhibitor Perfabloc SC (Fluka).

The method for extracting proteins from maize seeds for two-dimensional gel electrophoresis (2-D electrophoresis) was slightly modified from Consoli and Damerval (2001). Total proteins were extracted from 1 mg of seed powder by 75 µL of buffer containing 9.5 M urea, 2% v/v Triton X-100, 2% of IPG buffer pH range 3–10 (cat # 17-6000-87, GE), and 100 mM dithiothreitol (DTT). After incubating in a shaker at room temperature for 60 min, the sample was centrifuged for 10 min at 12,000 g in a bench-top micro-centrifuge, and the supernatant was carefully removed for the 2-D electrophoresis.

GFP detection by GFP antibody capture ELISA

The GFP content in maize seeds was quantified by using the enzyme-linked immunosorbent assay (ELISA). Goat anti-GFP antibody (Cat # 600-101-215, Rockland) was used at 2.5 µg/mL concentration, 50 µL per well. The ELISA plate (cat # 3590,

Corning, Inc.) was incubated for 1 hour at 37°C. The plate was then washed with 1 X with phosphate buffered saline with tween-20 or PBST [PBS: 0.01 M Na_2HPO_4, 0.003 M KH_2PO_4, 0.1 M NaCl (pH 7.2)]. The plate was then blocked with 5% dry milk in PBS for 1 hour at room temperature. After 1 X wash by PBST, 50 µL of the standard GFP (Cat # 4999-1000, BioVision) or the total water-soluble protein (TSP) from samples was added to the well and the plate was incubated at room temperature for 1 hour. After 3 X wash using PBST, 50 µL of chicken anti-GFP antibody (Cat # 600-901-215, Rockland) were diluted at 1:1000 with 1% dry milk in PBS and added to the well. The plate was incubated at 37°C for 1 hour. After 3 X wash by PBST, peroxidase conjugated rabbit anti-chicken antibody (Cat # 603-4302, Rockland) was diluted at 1:1000 with 1% dry milk in PBS and added to the well. After 3 X wash by PBST, 100 µL per well of peroxidase substrate (Cat # 52-00-02, KPL) was added. The plate was then read at 635 nm in a microplate reader (KC4, Biotek) after incubating at room temperature for 20 min.

Two-dimensional gel electrophoresis

Isoelectric focusing (IEF) was performed in a 7 cm long IPG ReadyStrip, pH 3-10 (cat # 163-2000, Bio-Rad). A sample of 125 µL was loaded onto the strip. Focusing was completed using the following steps—50 V for 12 hours, 500 V for 1 hour, 1000 V for 1 hour, and 8000 V for 8 hours at room temperature. The IEF was completed in the Protein Facility at Iowa State University. The second dimension and the Coomassie blue staining procedures were the same as described in *Protein extraction and Coomassie Brilliant Blue staining* in materials and methods, except the IPG well 12 % polyacrylamide SDS-PAGE gel (cat # 161-1391, Bio-Rad) was used instead of the regular well size gel. The

molecular weight marker was the Precision Plus Protein Standard Plugs (cat # 161-0378, Bio-Rad).

Statistical analysis

Statistical analyses of GFP expressions in T_3 transgenic maize were performed by utilizing the SAS v9.2 program (SAS Institute Inc., Cary, NC, USA). The general linear model (GLM) procedure was used. The whole data was considered as a randomized complete block design (RCBD).

Results and Discussion

Generation of transgenic maize line A222 (γRNAi)

The pX006 (Figure 1) is an RNAi (RNA interference) construct containing two identical copies of the gene coding sequence for the 27-kDa γ-zein—one in reverse direction and the other in forward direction. The two 27-kDa γ-zein coding sequences were separated by an intron 1355 bp long, containing a 703 bp chloramphenicol resistance gene (Cm^R) (Close and Rodriguez, 1982; Karimi et al., 2002) to facilitate the transcription forming a hairpin structure. The maize seed-specific promoter 27-kDa γ-zein promoter is used to drive the transgene cassette. The construct was introduced into the maize HiII germplasm by using an immature embryo via *Agrobacterium*-based methods (Frame et al., 2002).

The transgenic plants were brought to maturity in the greenhouse and crossed with the pollen donor of the inbred maize line B73. Initial transgene screens were performed on both calli and T_1 seeds by PCR. PCR indicated that all events were positive

for the transgenes of interest (data not shown). No significant difference was observed between the A222 line and wild-type maize during growth process. Fourteen independent events of transgenic maize line carrying pX006 (γ-zein down-regulation, designated as A222 or γRNAi) were generated, and a total of 3,403 kernels from 31 plants were harvested (Table 1).

Characterization of the γRNAi transgenic maize line A222

The opaque phenotype was observed from the mature seeds of the A222 line by using a white light transilluminator (Figure 2, Table 1). All 3,403 kernels from 31 plants representing 14 events were examined (Table 1). Opaque phenotype kernels were found in all events, except for A222-15. Some events, such as A222-1, A222-5, and A222-12, had low percentages of opaque seeds.

The opaque phenotype usually occurs when the zein content is reduced and the non-zein protein content is increased compensatorily or under stress (Crow and Kermicle, 2002; Lyznik and Tsai, 1989; Neuffer et al., 1997). To further confirm that the opaque phenotype of seeds resulted from the γ-zeins knock-down in these seeds, the total zein proteins were extracted from the A222 mature seeds and analyzed on the SDS-PAGE gel followed by Coomassie blue staining. Seeds from plants A222-2-1, A222-7-1, A222-10-2, and A222-13-1 were chosen for this analysis, since these four have many opaque phenotype kernels with bigger seed size. In Figure 3, the zein accumulation pattern is shown from six kernels randomly selected for each of the four events. The segregants with 27-kDa γ-zein and 16-kDa γ-zein knock-down pattern were observed for all four

events indicating the γ-zein RNAi construct (pX006) worked efficiently in transgenic maize line A222.

The γ-zein was knocked-down by RNAi in previous studies (Wu and Messing, 2010). The RNAi constructs used in Wu and Messing and our study were very similar. Both vectors contain a "reverse 27-kDa γ-zein – intron – forward 27-kDa γ-zein" RNAi structure and both are driven by the 27-kDa γ-zein promoter. The pattern for 27-kDa and 16-kDa γ-zein knock-down showed, by seed zein SDS-PAGE analysis in our study, was constant with observations with previous studies (Wu et al., 2010; Wu and Messing, 2010). The reduction of 22-kDa and 19-kDa α-zein was also observed from the seed zein SDS-PAGE analysis in the γRNAi QPM (quality protein maize) background (Wu et al., 2010), but not observed in our experiments.

The use of the γRNAi background maize to enhance the recombinant production under the control of the 27-kDa γ-zein promoter has not been addressed, yet. In the present study, we crossed the γRNAi line with the γGFP line and evaluated the recombinant protein (i.e. GFP) expression level in transgenic maize seeds.

Generation of hybrid γGFP/γRNAi

Two A222 (γRNAi) events (A222-2-1 and A222-13-1) were selected to cross with the P309 (γGFP, (Moeller et al., 2009)) line after zein characterization, since these two events have more bigger kernels and more healthy looking seeds than the others. The breeding strategy is shown in Figure 4. The A222 and P309 lines were crossed reciprocally. The T_2 generation was further advanced to T_3 by self-pollination. Only the double transgenic genotype (i.e., consisting of both γRNAi and γGFP) in T_2 was selected

for the parents of T_3. According to Mendel's genetic laws, 9 genotypes and 4 phenotypes should be generated in the T_3 generation. Two phenotypes (i.e., γRNAi with GFP and γRNAi without GFP) were further analyzed, and compared on the zein profiling and GFP expression levels.

GFP expression decreased in seeds with γRNAi background

To measure the GFP expression level in γGFP/γRNAi progenies, at least eight kernels were selected from each individual T_2 and T_3 progenies. First, genomic DNA was extracted from individually ground seed and PCR was performed to verify the genotype for each seed. GFP expression was quantified by ELISA for each individual GFP PCR positive seed. In addition, zein profiling was also completed on these seeds as shown in Figure 3. The seeds from the same plant were grouped into two phenotypes—GFP with γRNAi (G/Z) and GFP without γRNAi (G/z). The GFP expressions were compared between the G/Z and G/z kernels from the same plant.

In total, 40 T_2 progenies were generated from P309xA222-2-1 and P309xA222-13-1 crosses and analyzed by GFP ELISA and zein profiling. No significant difference of GFP level was observed between G/Z and G/z plants. The double transgenic plants were advanced to T_3 by self-pollination. A total of 43 T_3 progenies were generated from P309xA222-2-1. Of these, 20 progenies were verified as GFP positive by both PCR and ELISA. Among these 20 progenies, six were verified for γRNAi segregation (G/Z or G/z) by both PCR and zein profiling. The GFP-positive seeds were grouped by with or without γRNAi within each of the six progenies. The GFP expression data collected from individual kernel by ELISA are shown in Table 2. The mean GFP expression was

calculated by groups for each progeny. As can be seen in Table 3, five out of six progenies have higher mean GFP expression levels in seeds without γRNAi than in seeds with γRNAi. The GFP expression is significantly higher without γRNAi (G/z) than with γRNAi (G/z) in three T_3 progenies T3-1, T3-2, and T3-3, examined by the two-tail t-test (Table 3). The other three T_3 progenies show no significant difference between G/Z and G/z when using the t-test.

To further determine if the difference of the GFP expression level was significant, the statistical program SAS was used to analyze all samples. The data were analyzed by using the Randomized Complete Block Design (RCBD). Data obtained from all six progenies (Table 2) were considered as a block and the γRNAi was considered a treatment. The generalized linear model (GLM) procedure performed with SAS indicates the GFP expression difference between GFP lines without γRNAi and with γRNAi was significant (P <0.01). The two kernel types (i.e. G/Z and G/z) were the segragants from the same ear. Therefore, factors that cause the GFP expression difference in these two types of kernels, such as genetic background, growth and environmental conditions, and plants' individual differences, can be excluded. The only difference between G/Z and G/z kernels was the γRNAi. This result suggested the GFP expression under the control the 27-kDa γ-zein promoter decreased in maize lines when endogenous γ-zein genes were silenced.

This observation was opposite to what we originally anticipated. We hypothesized any recombinant and exogenous protein under γ-zein promoter would be enhanced in the γ-zein knock-down background, because reduced γ-zein content in total seed proteins

might generate a positive feedback to the γ-zein production mechanism; hence, promote the production of recombinant protein under the γ-zein promoter.

Recent studies about γ-zeins knock-down showed γRNAi slightly alters the protein body in morphology in which the zeins are stored, and the 19- and 22-kDa α-zeins are also reduced (Wu et al., 2010; Wu and Messing, 2010). The protein body is formed from the disulfide bonds cross-linked with β- and γ-zeins matrix first (Lending and Larkins, 1989; Lopes and Larkins, 1991). Then, the α- and δ-zeins are transferred into the β- and γ-zeins matrix during synthesis (Holding and Larkins, 2009; Kim et al., 2002; Lending and Larkins, 1989; Woo et al., 2001). The α-zeins are found in the protein body's core and the 27-kDa γ-zein is rich in the peripheral region (Holding et al., 2007). The γ and β-zeins coordinate the α- and δ-zeins and retain them in the ER (endoplasmic reticulum) for protein body formation (Bagga et al., 1997; Geli et al., 1994; Holding and Larkins, 2009; Kim et al., 2002). Previous studies also indicated both proteins would be secreted and degraded when α- and δ-zeins were expressed in petunia and tobacco (Bagga et al., 1997; Coleman et al., 1996; Williamson et al., 1988). Thus, the normal protein body structure relies on the organized accumulations of γ-, β-, α-, and δ-zeins, and the γ-zein plays a key role during protein body formation. When the γ-zein is knocked down, the protein body does not properly form, and the 19- and 22-kDa α-zeins are reduced (Wu et al., 2010; Wu and Messing, 2010).

In our study, the co-expressed γGFP levels in the γRNAi background were lower than their non-γRNAi segragants. Because limited cross events were evaluated, it is premature to conclude that recombinant protein expression driven by the γ-zein promoter will decrease in the γRNAi background. Research has shown the suppression of γ-zeins

leads to abnormal protein body formation. It is possible such disruption triggers a negative feedback (instead of a positive feedback hypothesized originally) to reduce any gene expression regulated by the γ-zein promoter (in our case, it is GFP). For plants, energy is wasted if production of a storage protein continues that cannot be stored in the proper place. On the other hand, when a major storage protein is reduced, other proteins compensatorily increase, as found in maize natural zein mutants *o2* and *fl2* (Coleman and Larkins, 1999), maize α-zein RNAi line (Wu and Messing, 2012b), and soybean β-conglycinin suppression line (Kinney et al., 2001). This makes sense that plants rebalance their protein content in seeds by producing more other proteins rather than enhancing the promoter for which the gene has been silenced.

One successful example of enhancing the foreign protein expression in crops by using a seed major protein silencing strategy has been completed in soybean; whereas, the GFP expression under glycinin promoter control was increased in the β-conglycinin silenced background (Schmidt and Herman, 2008). Since increased non-zein proteins were found in zein mutants and knock-down lines (Coleman and Larkins, 1999; Osborne et al., 1914; Wu and Messing, 2012b), it will be interesting to evaluate the GFP expression levels when GFP is driven by non-zein gene promoters in γRNAi maize background.

Generation of PBF overexpressing transgenic maize line A339 (oxPBF)

The pX012 construct carries a gene coding sequence for the native maize PBF (prolamin-box binding factor) driven by the maize ubiquitin promoter—a constitutive promoter. The construct was delivered into the maize HiII germplasm by using an

immature embryo via *Agrobacterium*-based methods. Sixteen transgenic lines carrying pX012 (oxPBF, designated as A339) were generated.

PBF is a transcription factor that regulates both 22-kDa α-zein and 27-kDa γ-zein (VicenteCarbajosa et al., 1997; Wu and Messing, 2012a). When PBF was suppressed by the PBF RNAi vector under the control of the 27-kDa γ-zein promoter, both 22-kDa α-zein and 27-kDa γ-zein were reduced, but other zeins were unaffected (Wu and Messing, 2012a). When PBF was over expressed, the expression of reporter genes (GUS or GFP) driven by the 27-Da γ-zein promoter was increased in immature maize seeds, embryos, calli and regenerated root tissues (Marzabal et al., 2008; Wu and Messing, 2012a).

Our intention is to co-express a recombinant protein (collagen) or marker protein (GFP) driven by the 27-Da γ-zein promoter in oxPBF background. Their expression levels will be compared with that in regular PBF background. We do not have these maize seeds available at this time. However, we have generated an oxPBF/γRNAi hybrid. The progenies of this hybrid have four genotypes——+/+; +/+, oxPBF/+; γRNAi/+, oxPBF/+; +/+, and +/+; +/γRNAi. The quantitative 2-D gel electrophoresis analysis was used to quantify the differences between proteins in the same position by measuring the density of the spot on the gel. Preliminary data indicated both 16- and 27-kDa γ-zeins disappeared in γRNAi genotypes (see black and white arrows in Figure 5). Our future plan is to quantify seed protein contents in the oxPBF background. If the up-regulated gene is discovered from these genotypes, the promoter can be used to enhance the recombinant protein expression in such backgrounds.

Summary and Future Work

In the present study, the γ-zeins (16- and 27-kDa γ-zein) in maize seed were successfully knocked down (Figure 3). The typical opaque phenotype was observed and correlated with the disappearance of 16- and 27-kDa γ-zeins. The γRNAi line was crossed with the 27-kDa γ-zein promoter-driving GFP (γGFP) line. Interestingly, the GFP levels in γGFP/γRNAi seeds appeared to be significantly lower than those for the non-γRNAi segregants. Because the suppression of γ-zeins production disrupts of protein body formation in seed, this, in turn, may impact negatively transgene expression under the control of the same γ-zein promoter. It is likely the plant chooses a more economical and energy savings way to have other proteins compensatorily increase rather than continue to produce a protein constantly suppressed. Therefore, it may be desirable to generate a GFP expressing maize line under a non-zein seed specific promoter (nzGFP). This nzGFP will be introduced into the γRNAi background. The nzGFP levels will then be compared with that of a non-γRNAi segregants.

A transcription factor for the 27-kDa γ-zein, PBF, was constitutively over-expressed in maize. We crossed the oxPBF line with the γRNAi line and generated progenies with four oxPBF and γRNAi genotype combinations. The next step is to examine how seed proteins change within these four different backgrounds by various quantitative proteomics. Difference gel electrophoresis (DIGE) will be used first. Seed proteins from different backgrounds would be labeled with different colors and then loaded on the same SDS-PAGE gel for electrophoresis. The protein content difference would be identified via the different intensity of the color from the same gel. Protein bands showing different intensities could be cut, collected, and further identified by MS.

We expect to evaluate the maize seed protein changes affected by either oxPBF or γRNAi, which could provide further information for maize seed nutrition improvement or enhancing the recombinant protein expression (i.e. use the increased protein promoter for expressing recombinant proteins in such background).

We also plan to cross the oxPBF line with the 27-kDa γ-zein promoter driving GFP line. Using this cross, we expect to achieve a higher GFP expression with over-expression of GFP. The GFP expression would be quantified by ELISA as described in the present study. That study should show whether it is worth using a co-expression of PBF strategy for expressing recombinant proteins driven by the 27-kDa γ-zein promoter.

Acknowledgments

We thank Dr. Lorena Moeller for offering the GFP transgenic line and sharing the GFP ELISA method, Margie Carter, Siquan Luo, Joel Nott, and Ann Perera in the Protein Facility and W.M. Keck Metabolomics Research Laboratory at Iowa State University for helpful discussions and suggestions, Fan Cao in the Department of Statistics for statistical analysis assistance, Dr. Susana Martin Ortigosa for helpful suggestions, and Angela Nguyen for technical assistance. This project is supported by the USDA Special grant IOW05228.

References

Bagga, S., Adams, H., Rodriguez, F., Kemp, J. and Sengupta Gopalan, C. (1997) Coexpression of the maize delta-zein and beta-zein genes results in stable

accumulation of delta-zein in endoplasmic reticulum-derived protein bodies formed by beta-zein. *Plant Cell* **9**, 1683-1696.

Barta, A., Sommergruber, K., Thompson, D., Hartmuth, K., Matzke, M. and Matzke, A. (1986) The expression of a nopaline synthase - human growth-hormone chimeric gene in transformed tobacco and sunflower callus-tissue. *Plant Molecular Biology* **6**, 347-357.

Basaran, P. and Rodrıguez-Cerezo, E. (2008) Plant molecular farming: opportunities and challenges. *Critical Reviews in Biotechnology* **28**, 153-172.

Borevitz, J., Xia, Y., Blount, J., Dixon, R. and Lamb, C. (2000) Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* **12**, 2383-2393.

Chikwamba, R., Cunnick, J., Hathaway, D., McMurray, J., Mason, H. and Wang, K. (2002) A functional antigen in a practical crop: LT-B producing maize protects mice against Escherichia coli heat labile enterotoxin (LT) and cholera toxin (CT). *Transgenic Research* **11**, 479-493.

Close, T. and Rodriguez, R. (1982) Construction and characterization of the chloramphenicol-resistance gene cartridge - a new approach to the transcriptional mapping of extrachromosomal elements. *Gene* **20**, 305-316.

Coleman, C., Herman, E., Takasaki, K. and Larkins, B. (1996) The maize gamma-zein sequesters alpha-zein and stabilizes its accumulation in protein bodies of transgenic tobacco endosperm. *Plant Cell* **8**, 2335-2345.

Coleman, C.E. and Larkins, B.A. (1999) 6. The Prolamins of Maize. *seed Proteins*, 109.

Consoli, L. and Damerval, C. (2001) Quantification of individual zein isoforms resolved by two-dimensional electrophoresis: Genetic variability in 45 maize inbred lines. *Electrophoresis* **22**, 2983-2989.

Crow, J. and Kermicle, J. (2002) Oliver Nelson and quality protein maize. *Genetics* **160**, 819-821.

Egelkrout, E., Rajan, V. and Howard, J. (2012) Overproduction of recombinant proteins in plants. *Plant Science* **184**, 83-101.

Esen, A. (1987) A proposed nomenclature for the alcohol-soluble proteins (zeins) of maize (Zea mays L.). *Journal of Cereal Science* **5**, 117-128.

Fischer, R. and Emans, N. (2000) Molecular farming of pharmaceutical proteins. *Transgenic Research* **9**, 279-299.

Frame, B.R., Shou, H.X., Chikwamba, R.K., Zhang, Z.Y., Xiang, C.B., Fonger, T.M., Pegg, S.E.K., Li, B.C., Nettleton, D.S., Pei, D.Q. and Wang, K. (2002) Agrobacterium tumefaciens-mediated transformation of maize embryos using a standard binary vector system. *Plant Physiology* **129**, 13-22.

Geli, M., Torrent, M. and Ludevid, D. (1994) 2 structural domains mediate 2 sequential events in gamma-zein targeting - protein endoplasmic-reticulum retention and protein body formation. *Plant Cell* **6**, 1911-1922.

Giddings, G. (2001) Transgenic plants as protein factories. *Current Opinion in Biotechnology* **12**, 450-454.

Holding, D. and Larkins, B. (2009) Zein storage proteins. *Molecular Genetic Approaches to Maize Improvement*: 269-286.

Holding, D., Otegui, M., Li, B., Meeley, R., Dam, T., Hunter, B., Jung, R. and Larkins, B. (2007) The maize floury1 gene encodes a novel endoplasmic reticulum protein involved in zein protein body formation. *Plant Cell* **19**, 2569-2582.

Hood, E. and Vicuna Requesens, D. (2011) Recombinant protein production in plants: challenges and solutions. *Methods In Molecular Biology: Recombinant Gene Expression Humana Press, New York*.

Hood, E., Witcher, D., Maddock, S., Meyer, T., Baszczynski, C., Bailey, M., Flynn, P., Register, J., Marshall, L., Bond, D., Kulisek, E., Kusnadi, A., Evangelista, R., Nikolov, Z., Wooge, C., Mehigh, R., Hernan, R., Kappel, W., Ritland, D., Li, C. and Howard, J. (1997) Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. *Molecular Breeding* **3**, 291-306.

Johnson, L.A., May, J.B. and White, P.J. (2003) Wet milling: the basis for corn biorefineries. In *Corn: Chemistry and Technology.* American Association of Cereal Chemists, 449-494.

Karimi, M., Inze, D. and Depicker, A. (2002) GATEWAY((TM)) vectors for Agrobacterium-mediated plant transformation. *Trends in Plant Science* **7**, 193-195.

Kermode, A. (2006) Plants as factories for production of biopharmaceutical and bioindustrial proteins: lessons from cell biology. *Canadian Journal of Botany-Revue Canadienne De Botanique* **84**, 679-694.

Kim, C., Woo, Y., Clore, A., Burnett, R., Carneiro, N. and Larkins, B. (2002) Zein protein interactions, rather than the asymmetric distribution of zein mRNAs on

endoplasmic reticulum membranes, influence protein body formation in maize endosperm. *Plant Cell* **14**, 655-672.

Kinney, A., Jung, R. and Herman, E. (2001) Cosuppression of the alpha subunits of beta-conglycinin in transgenic soybean seeds induces the formation of endoplasmic reticulum-derived protein bodies. *Plant Cell* **13**, 1165-1178.

Lee, J. and Forciniti, D. (2010) Purification of Human Antibodies from Transgenic Corn Using Aqueous Two-Phase Systems. *Biotechnology Progress* **26**, 159-167.

Lending, C. and Larkins, B. (1989) Changes in the zein composition of protein bodies during maize endosperm development. *Plant Cell* **1**, 1011-1023.

Liu, Q., Singh, S. and Green, A. (2002) High-oleic and high-stearic cottonseed oils: Nutritionally improved cooking oils developed using gene silencing. *Journal of the American College of Nutrition* **21**, 205S-211S.

Lopes, M. and Larkins, B. (1991) Gamma-zein content is related to endosperm modification in quality protein maize. *Crop Science* **31**, 1655-1662.

Lyznik, L. and Tsai, C. (1989) Protein-synthesis in endosperm cell-cultures of maize (zea-mays L.). *Plant Science* **63**, 105-114.

Ma, J., Drake, P. and Christou, P. (2003) The production of recombinant pharmaceutical proteins in plants. *Nature Reviews Genetics* **4**, 794-805.

Marzabal, P., Gas, E., Fontanet, P., Vicente-Carbajosa, J., Torrent, M. and Ludevid, M. (2008) The maize Dof protein PBF activates transcription of gamma-zein during maize seed development. *Plant Molecular Biology* **67**, 441-454.

Mehrtens, F., Kranz, H., Bednarek, P. and Weisshaar, B. (2005) The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiology* **138**, 1083-1096.

Mena, M., Vicente-Carbajosa, J., Schmidt, R. and Carbonero, P. (1998) An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm. *Plant Journal* **16**, 53-62.

Moeller, L., Gan, Q. and Wang, K. (2009) A bacterial signal peptide is functional in plants and directs proteins to the secretory pathway. *Journal of Experimental Botany* **60**, 3337-3352.

Muntz, K. (1998) Deposition of storage proteins. *Plant Molecular Biology* **38**, 77-99.

Murray, M.G. and Thompson, W.F. (1980) Rapid isolation of high molecular-weight plant DNA. *Nucleic Acids Research* **8**, 4321-4325.

Neuffer, M.G., Coe, E.H. and Wessler, S.R. (1997) *Mutants of Maize*. Plainview, NY: Cold Spring Harbor Laboratory Press.

Obembe, O., Popoola, J., Leelavathi, S. and Reddy, S. (2011) Advances in plant molecular farming. *Biotechnology Advances* **29**, 210-222.

Osborne, T., Mendel, L., Ferry, E. and Wakeman, A. (1914) Nutritive properties of proteins of the maize kernel. *Journal of Biological Chemistry* **18**, 1-16.

Paraman, I., Moeller, L., Scott, M., Wang, K., Glatz, C. and Johnson, L. (2010) Utilizing protein-lean coproducts from corn containing recombinant pharmaceutical proteins for ethanol production. *Journal of Agricultural and Food Chemistry* **58**, 10419-10425.

Ramessar, K., Sabalza, M., Capell, T. and Christou, P. (2008) Maize plants: An ideal production platform for effective and safe molecular pharming. *Plant Science* **174**, 409-419.

Schmidt, M. and Herman, E. (2008) Proteome rebalancing in soybean seeds can be exploited to enhance foreign protein accumulation. *Plant Biotechnology Journal* **6**, 832-842.

Shiu, S.H., Shih, M.C. and Li, W.H. (2005) Transcription factor families have much higher expansion rates in plants than in animals. *Plant Physiology* **139**, 18-26.

Sodek, L. and Wilson, C. (1971) Amino acid composition of proteins isolated from normal, opaque-2, and floury-2 corn endosperms by a modified osborne procedure. *Journal of Agricultural and Food Chemistry* **19**, 1144-&.

Stoger, E., Ma, J., Fischer, R. and Christou, P. (2005) Sowing the seeds of success: Pharmaceutical proteins from plants. *Current Opinion in Biotechnology* **16**, 167-173.

Stracke, R., Ishihara, H., Barsch, G., Mehrtens, F., Niehaus, K. and Weisshaar, B. (2007) Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the Arabidopsis thaliana seedling. *Plant Journal* **50**, 660-677.

Ueda, T. and Messing, J. (1991) A homologous expression system for cloned zein genes. *Theoretical and Applied Genetics* **82**, 93-100.

VicenteCarbajosa, J., Moose, S., Parsons, R. and Schmidt, R. (1997) A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic

leucine zipper transcriptional activator Opaque2. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 7685-7690.

Williamson, J., Galili, G., Larkins, B. and Gelvin, S. (1988) The synthesis of a 19-kilodalton zein protein in transgenic petunia plants. *Plant Physiology* **88**, 1002-1007.

Wolt, J. and Karaman, S. (2007) Estimated environmental loads of alpha-amylase from transgenic high-amylase maize. *Biomass & Bioenergy* **31**, 831-835.

Woo, Y., Hu, D., Larkins, B. and Jung, R. (2001) Genomics analysis of genes expressed in maize endosperm identifies novel seed proteins and clarifies patterns of zein gene expression. *Plant Cell* **13**, 2297-2317.

Wu, Y., Holding, D. and Messing, J. (2010) gamma-Zeins are essential for endosperm modification in quality protein maize. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 12810-12815.

Wu, Y. and Messing, J. (2010) RNA interference-mediated change in protein body morphology and seed opacity through loss of different zein proteins. *Plant Physiology* **153**, 337-347.

Wu, Y. and Messing, J. (2012a) Rapid Divergence of Prolamin Gene Promoters of Maize after Gene Amplification and Dispersal. *Genetics*.

Wu, Y. and Messing, J. (2012b) RNA Interference Can Rebalance the Nitrogen Sink of Maize Seeds without Losing Hard Endosperm. *Plos One* **7**.

Xu, J., Dolan, M.C., Medrano, G., Cramer, C.L. and Weathers, P.J. (2011a) Green Factory: Plants as bioproduction platforms for recombinant proteins. *Biotechnology Advances* **30**:1171-1184

Xu, J. and Messing, J. (2008) Organization of the prolamin gene family provides insight into the evolution of the maize genome and gene duplications in grass species. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 14330-14335.

Xu, X., Gan, Q., Clough, R., Pappu, K., Howard, J., Baez, J. and Wang, K. (2011b) Hydroxylation of recombinant human collagen type I alpha 1 in transgenic maize co-expressed with a recombinant human prolyl 4-hydroxylase. *Bmc Biotechnology* **11**: 69.

Zarkadas, C., Hamilton, R., Yu, Z., Choi, V., Khanizadeh, S., Rose, N. and Pattison, P. (2000) Assessment of the protein quality of 15 new northern adapted cultivars of quality protein maize using amino acid analysis. *Journal of Agricultural and Food Chemistry* **48**, 5351-5361.

Zhong, Q., Xu, L., Zhang, C. and Glatz, C. (2007) Purification of recombinant aprotinin from transgenic corn germ fraction using ion exchange and hydrophobic interaction chromatography. *Applied Microbiology and Biotechnology* **76**, 607-613.

Table 1 Generation of transgenic maize line A222 (γRNAi)

	PTF ID	Event	Plant	# total seed	*# opaque seed	% op/total	**Zein Coomassie
1	A222	1	2	141	5	3.5%	
	A222	1	3	184	0	0.0%	
	A222	1	4	247	5	2.0%	
		subtotal		572	10	1.7%	
2	A222	2	1	81	70	86.4%	See Figure3
	A222	2	3	38	33	86.8%	
	A222	2	4	44	10	22.7%	
		subtotal		163	113	69.3%	
3	A222	5	1	238	5	2.1%	
4	A222	6	1	28	13	46.4%	
	A222	6	3	39	8	20.5%	
	A222	6	4	103	10	9.7%	
		subtotal		408	36	8.8%	
5	A222	7	1	25	22	88.0%	See Figure3
6	A222	10	2	62	60	96.8%	See Figure3
	A222	10	3	56	52	92.9%	
	A222	10	4	6	6	100.0%	
		subtotal		149	140	94.0%	
7	A222	12	1	222	0	0.0%	
	A222	12	3	202	2	1.0%	
	A222	12	4	229	0	0.0%	
		subtotal		653	2	0.3%	
8	A222	13	1	205	59	28.8%	See Figure3
9	A222	15	2	271	0	0.0%	
10	A222	17	2	153	0	0.0%	
	A222	17	3	62	0	0.0%	
	A222	17	4	220	2	0.9%	
		subtotal		911	61	6.7%	
11	A222	18	4	3	3	100.0%	
12	A222	19	1	43	39	90.7%	
	A222	19	2	6	6	100.0%	
	A222	19	4	9	9	100.0%	
		subtotal		61	57	93.4%	
13	A222	20	1	93	79	84.9%	
	A222	20	2	159	10	6.3%	
		subtotal		252	89	35.3%	
14	A222	22	1	2	2	100.0%	
	A222	22	3	29	29	100.0%	
	A222	22	4	203	24	11.8%	
		subtotal		234	55	23.5%	
		total		3403	563	16.5%	

*Determined by using a white light transilluminator.

**Coomassie Brilliant Blue staining of total zein extraction from seeds.

Table 2. T_3 seed ^GFP expression with (G/Z) or without γ-zein knock-down (G/z)

T3-1		T3-2		T3-3		T3-4		T3-5		T3-6	
G/Z	G/z	G/Z	G/z	G/Z	G/z	G/Z	G/z	G/Z	G/z	G/Z	G/z
0.05	0.49	0.10	0.22	0.21	0.66	0.14	0.09	0.17	0.70	0.16	0.63
0.09	0.84	0.10	0.74	0.26	0.85	0.21	0.40	0.18	0.94	0.19	0.90
0.10	0.90	0.28	0.97	0.35	3.39	0.30	0.44	0.19	0.97	0.23	0.95
0.11	1.03	0.39	1.20	0.38	4.05	0.30	0.49	0.19	1.12	0.43	3.60
0.12	1.20	0.43	1.26	0.38	4.58	0.36	0.58	0.20	1.17	0.43	6.62
0.28	0.89	0.46	0.88	0.40		0.44	1.13	0.23	1.37	0.44	
0.33	0.27	0.47	0.42	0.41		0.45	1.23	0.28	1.49	0.58	
0.34		0.48		0.43		0.55		0.29	3.07	0.84	
0.49		0.34		0.50		0.61		0.30	3.63	0.99	
0.58		0.16		0.56		0.64		0.32		1.08	
0.82				0.58		0.79		0.40		1.50	
0.30				0.58		1.04		0.42		1.59	
0.24				0.65		1.32		0.44		1.77	
				0.67		1.37		0.48		2.88	
				0.68		2.33		0.48		3.23	
				0.82		2.38		0.50		3.80	
				0.84		2.54		0.56		3.86	
				0.86				0.56		4.33	
				0.92				0.59		4.66	
				0.99				0.77			
				1.02				0.78			
				1.03				1.01			
				1.11				1.13			
				1.15				1.76			
				1.24				1.83			
				1.39				3.25			
				1.47				3.38			
				1.57				4.96			
				1.70							
				2.14							
				2.30							
				2.92							
				3.49							
				4.92							
				5.76							

^GFP expression levels are ug per mg of total protein

Table 3. GFP expression in T_3 maize seed (P309xA222-2-1) with or without γ-zein knock-down

		GFP with γRNAi (G/Z)		GFP w/o γRNAi (G/z)		two-tail t-test
Events	Total kernel # analyzed	kernel #	^Mean GFP expression	kernel #	*Mean GFP expression	P-value
T3-1	20	13	0.30	7	0.80	*0.000611
T3-2	17	10	0.32	7	0.81	*0.002094
T3-3	40	35	1.28	5	2.71	*0.030800
T3-4	24	17	0.93	7	0.62	0.351047
T3-5	37	28	0.92	9	1.61	0.119340
T3-6	24	19	1.74	5	2.54	0.379165

^GFP expression levels are ug per mg of total protein.

*Significance in t-test, P-value < 0.05.

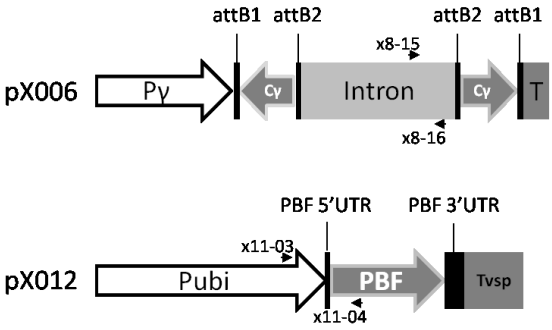

Figure 1. Two constructs used in this study. In construct pX006: Pγ, 27-kDa γ-zein promoter; attB1 and attBt, recombination sites for Gateway cloning; Cγ, 27-kDa γ-zein gene coding sequence; T, 35S terminator; ; x8-15 and x8-16, primers for verifying the presence of γRNAi transgene cassette in pX006. In construct pX012: Pubi, maize ubiquitin promoter; PBF, maize prolamin-box binding factor coding sequence; PBF 5'UTR, PBF 5' untranslated region; PBF 3'UTR, PBF 3' untranslated region; Tvsp, the soybean vegetative storage protein gene terminator; x11-03 and x11-04, primers for verifying the presence of PBF transgene cassette in pX012.

Figure 2. Opaque phenotype of maize seeds. Opaque phenotype is observed in γ-zein reduced transgenic maize seed A222-9-2. WT is wild type. Photographs were taken with transmitted light.

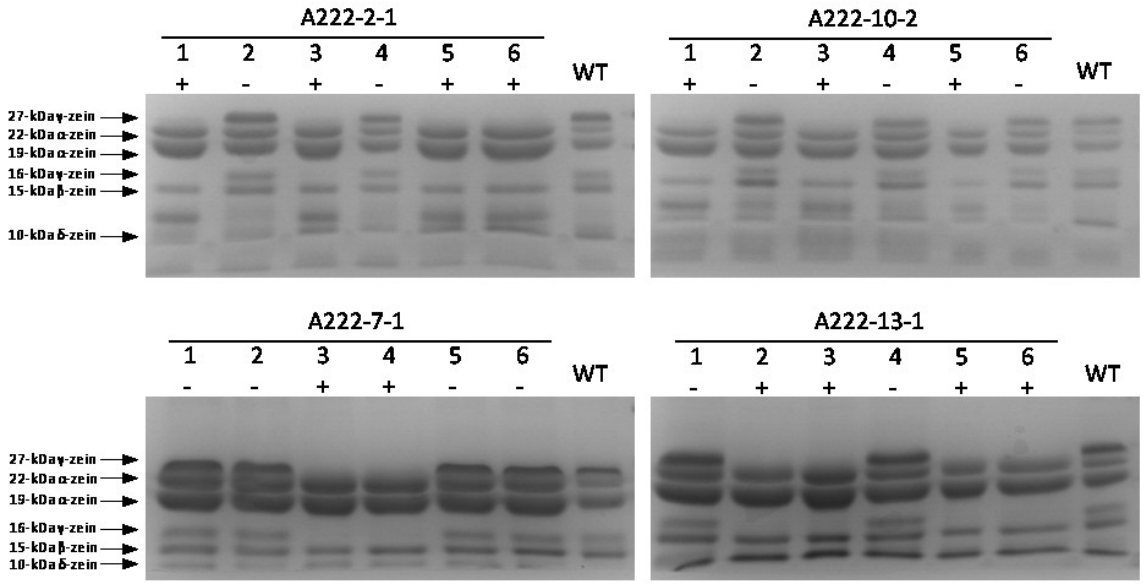

Figure 3. Maize seed zein accumulation pattern of four events from T_1 A222 (γRNAi). Total zein proteins were extracted from six seeds from each event and detected by SDS-PAGE followed by the Coomassie staining. Segregants with 27-kDa γ-zein and 16-kDa γ-zein knock-down pattern are observed in all four events and labeled as "+" on the top of the lane. The negative segregants are labeled as "-". The WT is wild type B73.

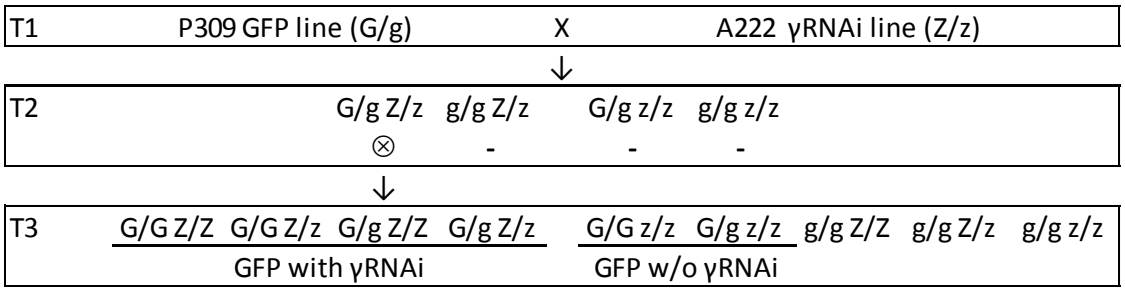

Figure 4. Breeding schemes of double transgenic plants. The T1 GFP was crossed with γRNAi. Four genotypes were generated in T2 and only one (G/g Z/z) was selected to advance to T3 by self- pollination. γRNAi, γ-zein knock-down; GFP, green fluorescent protein over expression; Z, dominant transgene of γ-zein RNAi; G, dominant transgene of GFP.

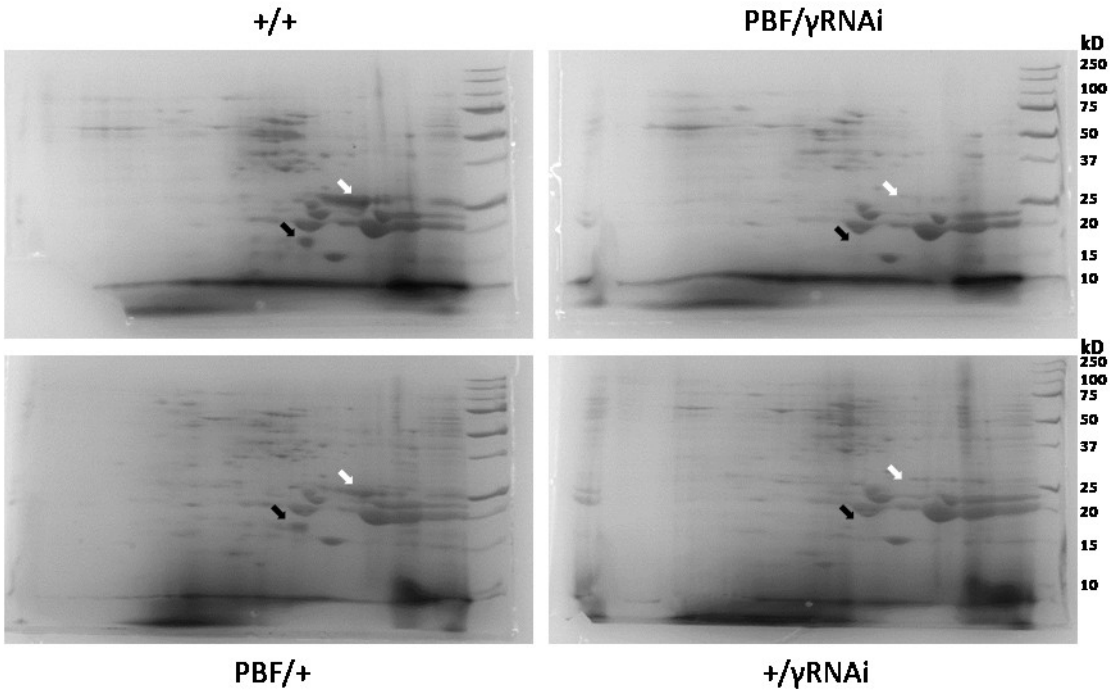

Figure 5. Two-D patterns of 4 genotypes T1 of A339 (PBF) X A222 (γRNAi). γRNAi, γ-zein knock-down; PBF, prolamin-box-binding factor over expression. Dark arrows indicate the 16-kDa γ-zein, and white arrows indicate the 27-kDa γ-zein, which are all absent in γRNAi segregates.

CHAPTER 7: GENERAL CONCLUSIONS

In this dissertation, we proposed that maize can be used to safely produce high yielding and adequately modified recombinant proteins in seeds. Through Chapters 3-6, I presented research activities aimed at demonstrating this proposed approach: 1) produce hydroxylated recombinant human collagen type I α 1 (rCIα1) by co-expression of recombinant human prolyl 4-hydroxylase in maize seeds (Chapter 3); 2) generate a small biosurfactant peptide GAM1 in maize seeds with fusion to the red fluorescent protein (RFP) or a maize seed storage protein 22-kDa α-zein (Chapter 4); 3) establish transformation of the tissue culture-amenable male-sterile maize and transfer a vaccine gene, the *E-coli* heat-labile enterotoxin B subunit (LT-B), from male-fertile to sterile maize lines for open-field release (Chapter 5); and 4) evaluate the green fluorescent protein (GFP) expression in maize seeds with either silencing of a seed storage protein, the 27 kDa γ-zein or overexpression of a transcription factor prolamin box-binding factor (PBF) (Chapter 6).

In Chapter 3, the rCIα1 was successfully expressed in high percentages of hydroxylated proline (Hyp) in maize seeds under the control of the seed-specific promoter maize globulin-1, when co-expressed with a recombinant human prolyl 4-hydroxylase (rP4H). High-resolution mass spectrometry (HRMS) analysis revealed that the Hyp ratio of maize-derived rCIα1with co-expression of rP4H (rCIα1-OH) is 18.11%, higher than the rCIα1 without co-expression of the rP4H (5.89%) and even higher than the native human CIα1 (14.95%). The increased Hyp content was correlated with distinctly enhanced thermal stability of rCIα1 compared to the non-hydroxylated rCIα1. We concluded maize is able to produce recombinant proteins with mammalian-like post-translational modifications, which may be necessary for their use as pharmaceutical and industrial productions.

In Chapter 4 the protein fusion strategy was presented to produce a 22-amino acid biosurfactant peptide GAM1 in maize seeds by fusing to the RFP or the 22-kDa α-zein. Maize endogenous seed-specific promoters of 27-kDa γ-zein and 22-kDa α-zein were used to lead the expression of RFP::GAM1 and α-zein::GAM1, respectively. Correctly fused sequences were confirmed by PCR and DNA sequencing for both constructs. Transgenic maize seeds carrying RFP::GAM1 displayed red fluorescence and were further confirmed as RFP positive in Western Blot analysis. Various analysis methods including in-gel digestion/peptide mass fingerprinting, high-performance liquid chromatography (HPLC), and MS have been used to detect the presence of the GAM1 in both RFP::GAM1 and Zein::GAM1 fusion lines. Due to low levels of fusion protein production and ineffective protein concentration methods, the verification of GAM1 was not successful. Future work should include reverse transcription PCR (RT-PCR) to determine the presence of the full length of fusion transcripts in seeds. The fusion and gene construction strategies will also need to be reevaluated.

In Chapter 5, two strategies were described for the safe production of maize-derived pharmaceutical proteins in open field by using a male-sterile transgenic maize line to avoid pollen contamination. The first strategy was to directly introduce the transgene into transformable male-sterile maize. Transgene cassette maize ubiquitin promoter driving both *bar* and *gus*-intron was used for biolistic-mediated (gene-gun) transformation method, and 35S promoter driving both *bar* and *gus*-intron cassette was used for *Agrobacterium*-based transformation methods. GUS assay was performed to verify the expression of *gus* gene in transgenic maize callus and leaves in the R_1 progeny. This is the first report of successful transformation of male-sterile maize by biolistic- or *Agrobacterium*-based methods. The

second strategy was to introduce the transgene from a male-fertile transgenic line to the male-sterile maize line via conventional breeding. The B-subunit protein of *Escherichia coli* heat labile enterotoxin (LT-B), a vaccine gene, was used as a pharmaceutical product model in this strategy. After six seasons of breeding, we were able to obtain 100% transgenic male-sterile maize seeds by using non-transgenic pollen donor in an open field. We concluded that maize-derived pharmaceutical proteins can be produced in the open field using male-sterile maize lines.

In Chapter 6, maize seed protein γ-zeins (16- and 27-kDa γ-zein) were successfully knocked down via RNAi (RNA interference). By introgression of the 27-kDa γ-zein promoter driving GFP (γGFP) into the γ-zeins knock-down (γRNAi) maize line, we examined whether the GFP expression could be enhanced in γ-zein reduced maize background. The GFP expression in γGFP/γRNAi seeds was significantly lower than that of in the non-γRNAi segregants. This unexpected observation may have been due to the fact that both γ-zein knock-down and GFP expression were controlled by the same 27-kDa γ-zein promoter. Since the reduction of γ-zeins disrupts protein body formation in maize seeds, it may in turn negatively regulate the foreign gene expression driven by the same γ-zein promoter. It may be an energy-saving strategy for plants so other proteins compensatorily increased instead of continuing to produce a protein that is constantly suppressed. Thus, it would be desirable to use a non-zein seed specific promoter to lead the expression of the GFP (nzGFP) in the γRNAi background. The expression levels will then be compared between the γRNAi and non-γRNAi segregants.

Another attempt described in Chapter 6 was to enhance the expression of recombinant proteins in transgenic maize by overexpressing a transcription factor for the 27-kDa γ-zein,

prolamin-box binding factor (PBF). We hypothesized that the expression of the 27-kDa γ-zein promoter driving recombinant protein is enhanced by over-expressing PBF in maize seeds. The PBF over-expression transgenic maize line (oxPBF) was successfully generated. Future work should include crossing the oxPBF line with the 27-kDa γ-zein promoter driving GFP line. GFP expression should be quantified by ELISA as described in this study. This study would provide further information of whether it is a good strategy to co-express the PBF to enhance recombinant proteins driven by the 27-kDa γ-zein promoter in maize seeds.

The studies presented in this dissertation demonstrate strategies to overcome challenges encountered in using transgenic maize to produce recombinant proteins. While maize has the potential to produce recombinant proteins with mammalian-like post-translation modifications and can be produced safely in an open field, it remains challenging to produce a high quantity in seeds for industrial purposes. The materials and preliminary data related to the regulatory mechanism of maize seed protein generated from this work can be used to further improve recombinant protein expression in seeds.